

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**



Access DB# 99955

## RESEARCH REQUEST FORM

JUL 29 2003  
Scientific and Technical Information Center

Requester's Full Name: K-C. SRIVASTAVA Examiner #: 77964 Date: 07/29/2003  
 Art Unit: 1651 Phone Number 301-605-1196 Serial Number: 10/087,195  
 Mail Box and Bldg/Room Location: TP504 Results Format Preferred (circle): PAPER DISK E-MAIL  
Room # CM-11A12

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched.  
 Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: TREATMENT OF ALLERGIC RHINITISInventors (please provide full names): PATRICIA ANNE NUTALL &  
GUIDO CHRISTIAN PAESENEarliest Priority Filing Date: 09/01/1999

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search for Protein sequence

AA 182 MS-HBPI = SEQ ID # 8

AA 172 FS-HBPI = SEQ ID # 6

AA 171 ES-HBPI = SEQ ID # 7

D. RET 6 = ?

TICK = RHIPKEPHALUS <sup>SINUS</sup> EVERTI  
<sup>SANGUINEUS</sup>And all terms ~~related~~ with ALLERGIC  
RHINITIS  
+ TREATMENT

Due 09/15/2003

+ above SEQ ID'S  
HISTACALIN PROTEIN" ECTOPARASITE (TICK  
OR - AMBLYOMMA <sup>SP</sup> ~~SP~~ <sup>AMERICAN</sup>  
BLOOD FEEDING" SEASONAL OR PERENNIAL  
OR CHRONIC OR ACUTE  
+ ALLERGIC RHINITIS

## STAFF USE ONLY

Searcher: Hanby Type of Search: NA Sequence (#) STN \$182  
 Searcher Phone #: 301-605-1196 AA Sequence (#) 34 Dialog 02  
 Searcher Location: TP504 Structure (#) Questel/Orbit 02  
 Date Searcher Picked Up: 8/1 Bibliographic Dr. Link 02  
 Date Completed: 8/4 Litigation Lexis/Nexis 02  
 Searcher Prep & Review Time: SEQ 20, STN 30 Fulltext Sequence Systems 02  
 Clerical Prep Time: 0 Patent Family WWW/Internet 02  
 Online Time: SEQ 20, STN 35 Other Other (specify) 02

**THIS PAGE BLANK (USPTO)**



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 99955**

**TO: Kailash C Srivastava**  
**Location: cm-1/11a12/11b01**  
**Art Unit : 1651**  
**Monday, August 04, 2003**

**Case Serial Number: 10/087195**

**From : Susan Hanley**  
**Location: Biotech-Chem Library**  
**CM1 6B05**  
**Phone: 305-4053**

**susan.hanley@uspto.gov**

### **Search Notes**

**THIS PAGE BLANK (USPTO)**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor  
308-4258; CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem/Library CM1 - Circ. Desk



**THIS PAGE BLANK (USPTO)**

# Registry Records for FSHBP1 & 2 & MSHBP1 & 2

D. RET 6

SRIVASTAVA 10/087,195

=> d sqide 1

L6 ANSWER 1 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN **MSHBP1**  
 RN 329085-10-1 REGISTRY  
 CN 3: PN: W00116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)  
 OTHER NAMES:  
 CN 3: PN: W00115719 TABLE: 1 unclaimed protein  
 FS PROTEIN SEQUENCE  
 SQL 182

## PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
Not Given	W02001015719
	unclaimed
	TABLE 1
	W02001016164
	unclaimed
	TABLE 1

SEQ 1 NPTWANEAKL GSYQDAWKS QDQNKRYL AQATQTTDGV WGEEFTCVSV  
 51 TAEKIGKKKL NATILYKNKH LTDLKESHET ITWKAYDYT TENGIKYEYQ  
 101 GTRTQTFEDV FVFSYKNC D VIFVPKERGS DEGDYELWVS EDKIDKIPDC  
 151 CKFTMAYFAQ QKEKTVRNVY TDSSCKPAPA QN  
 MF Unspecified  
 CI MAN  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL  
 2 REFERENCES IN FILE CA (1947 TO DATE)  
 2 REFERENCES IN FILE CAPLUS (1947 TO DATE)

=> d sqide 2

L6 ANSWER 2 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN **FSHBP2**  
 RN 329085-09-8 REGISTRY  
 CN 2: PN: W00116164 TABLE: 1 unclaimed sequence (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE  
 SQL 171

## PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
Not Given	W02001016164
	unclaimed
	TABLE 1

SEQ 1 NQPDWADEAA NGAHQDAWKS LKADVENVYY MVKATYKNDP VWGNDFTCVG  
 51 VMANDVNEDE KSIQAEFLM NNADTNMQFA TEKVTAVKMY GYNRENAFRY  
 101 ETEDGQVFTD VIAYSDDNCD VIYVPGTDGN EEGYELWTTD YDNIPANCLN  
 151 KFNEYAVGRE TRDVFTSACL E  
 MF Unspecified  
 CI MAN  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

**THIS PAGE BLANK (USPTO)**



1 REFERENCES IN FILE CA (1947 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1947 TO DATE)

=> d sqide 3

L6 ANSWER 3 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 329085-08-7 REGISTRY  
CN 1: PN: WO0116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN 1: PN: WO0115719 TABLE: 1 unclaimed protein  
FS PROTEIN SEQUENCE  
SQL 172

FS-4BPI

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
Not Given	WO2001015719
	unclaimed
	TABLE 1
	WO2001016164
	unclaimed
	TABLE 1

SEQ 1 DKPVWADEAA NGEHQDAWKH LQKLVEENYD LIKATYKNDP VWGNDFTCVG  
51 TAAQNLNEDE KNVEAWFMFM NNADTVYQHT FEKATPDKMY GYNKENAITY  
101 QTEDGQVLTD VLAFSDDNCY VIYALGPDGS GAGYELWATD YTDVPASCLE  
151 KFNEYAAGLP VRDVYTSACL PE

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL  
2 REFERENCES IN FILE CA (1947 TO DATE)  
2 REFERENCES IN FILE CAPLUS (1947 TO DATE)

**THIS PAGE BLANK (USPTO)**

=> d sqide 128

D. RET6

L28 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 200220-35-5 REGISTRY  
CN Protein D.RET6 (Dermacenter reticularis histamine-binding) (9CI) (CA  
INDEX NAME)  
FS PROTEIN SEQUENCE  
SQL 209

SEQ 1 MKMQVLLLT FVSAALATQA ETTSKAGEN PLWAHEELG KYQDAWSID  
51 QGVSVTVVLA KTTYENDTGS WGSQFKCLQV QEIERKEEDY TVTSVFTFRN  
101 ASSPIKYYNV TETVKAVFY GYKNIRNAIE YQVGGGLNIT DTLIFTDGEL  
151 CDVFYVPNAD QGCELWVKKS HYKHVPDYCT FVFNVCAKD RKTYDIFNEE  
201 CVYNGEPWL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS  
2 REFERENCES IN FILE CA (1947 TO DATE)  
2 REFERENCES IN FILE CAPLUS (1947 TO DATE)

**THIS PAGE BLANK (USPTO)**

This is a review on the HBP subject  
but there are no Registry #s associated  
w/ it

SRIVASTAVA 10/087,195

=> d ibib abs hitstr ind

L31 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2000:799221 HCAPLUS  
DOCUMENT NUMBER: 134:55073  
TITLE: Vector-host interactions in disease transmission  
AUTHOR(S): Nuttall, P. A.; Paesen, G. C.; Lawrie, C. H.; Wang, H.  
CORPORATE SOURCE: CEH Institute of Virology and Environmental  
Microbiology, Oxford, OX1 3SR, UK  
SOURCE: Journal of Molecular Microbiology and Biotechnology  
(2000), 2(4), 381-386  
CODEN: JMMBFF; ISSN: 1464-1801  
PUBLISHER: Horizon Scientific Press  
DOCUMENT TYPE: Journal; General Review  
LANGUAGE: English

AB A review with 56 refs. Tick-borne spirochetes include borreliae that cause Lyme disease and relapsing fever in humans. They survive in a triangle of parasitic interactions between the spirochete and its vertebrate host, the spirochete and its tick vector, and the host and the tick. Until recently, the significance of vector-host interactions in the transmission of arthropod-borne disease agents has been overlooked. However, there is now compelling evidence that the pharmacol. activity of tick saliva can have a profound effect on pathogen transmission both from infected tick to uninfected host, and from infected host to uninfected tick. The salivary glands of ticks provide a pharmacopoeia of anti-inflammatory, anti-hemostatic and anti-immune mols. These include bioactive proteins that control histamine, bind Igs, and inhibit the alternative complement cascade. The effect of these mols. is to provide a privileged site at the tick-host interface in which borreliae and other tick-borne pathogens are sheltered from the normal innate and acquired host immune mechanisms that combat infections. Understanding the key events at the tick vector-host interface, that promote spirochete infection and transmission, will provide a better understanding of the epidemiol. and ecol. of these important human pathogens.

CC 15-0 (Immunochemistry)

Section cross-reference(s): 12

ST review immunomodulator tick Borrelia transmission

IT Proteins, specific or class

RL: BOC (Biological occurrence); BSU (Biological study, unclassified);  
BIOL (Biological study); OCCU (Occurrence)

(Ig-binding proteins; tick vector-host interaction and Borrelia  
transmission in relation to salivary expression of)

IT Complement

(activation; tick vector-host interaction and Borrelia transmission in  
relation to salivary expression of inhibitors of)

IT Proteins, specific or class

RL: BOC (Biological occurrence); BSU (Biological study, unclassified);  
BIOL (Biological study); OCCU (Occurrence)

(histacalins; tick vector-host interaction and Borrelia  
transmission in relation to salivary expression of)

IT Borrelia

Ixodes

Lyme disease

Saliva

(salivary immunomodulators in tick vector-host interaction and Borrelia  
transmission)

REFERENCE COUNT: 56 THERE ARE 56 CITED REFERENCES AVAILABLE FOR THIS  
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

**THIS PAGE BLANK (USPTO)**

=> d que 132

L5 4 SEA FILE=REGISTRY ABB=ON PLU=ON (329085-08-7/BI OR 329085-09-8/BI OR 329085-10-1/BI OR 51-45-6/BI)  
L6 3 SEA FILE=REGISTRY ABB=ON PLU=ON L5 AND PROTEIN/FS  
L9 3 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTACALIN?  
L10 2 SEA FILE=HCAPLUS ABB=ON PLU=ON L6  
L11 4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP1  
L13 4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP2  
L14 4 SEA FILE=HCAPLUS ABB=ON PLU=ON MS-HBP1  
L15 4 SEA FILE=HCAPLUS ABB=ON PLU=ON "D.RET6"  
L17 18 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN  
L18 4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-1  
L19 4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-2  
L24 6 SEA FILE=HCAPLUS ABB=ON PLU=ON (L9 OR L10) OR L11 OR (L13 OR L14 OR L15) OR (L18 OR L19)

*citations  
w/ Reg #'s*

~~L32 14 SEA FILE=HCAPLUS ABB=ON PLU=ON L17 NOT L24~~

*← cites on HBP proteins  
but don't have the  
Registry #'s for  
D.RET6, SFHBP1, 2 or  
MS HBP1*

**THIS PAGE BLANK (USPTO)**



=&gt; d ibib abs hitstr 1-14 ind

L32 ANSWER 1 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:492892 HCAPLUS

TITLE: RNA interference in ticks: a study using  
**histamine binding protein**dsRNA in the female tick *Amblyomma americanum*AUTHOR(S): Aljamali, M. N.; Bior, A. D.; Sauer, J. R.; Essenberg,  
R. C.CORPORATE SOURCE: Department Biochemistry and Molecular Biology,  
Oklahoma State University, Stillwater, OK, 74078, USA

SOURCE: Insect Molecular Biology (2003), 12(3), 299-305

CODEN: IMBIE3; ISSN: 0962-1075

PUBLISHER: Blackwell Publishing Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB RNA interference (RNAi), a gene silencing process, has been recently exploited to det. gene function by degrading specific mRNAs in several eukaryotic organisms. We constructed a double stranded RNA (dsRNA) from a previously cloned putative *Amblyomma americanum* **histamine binding protein** (HBP) to test the significance of using this methodol. in the assessment of the function and importance of gene products in ectoparasitic ticks. The female salivary glands incubated in vitro with HBP dsRNA had a significantly lower histamine binding ability. In addn., the injection of HBP dsRNA into the unfed females led both to a reduced histamine binding ability in the isolated salivary glands and to an aberrant tick feeding pattern or host response. Mol. data demonstrated less expression of the HBP mRNA in the RNAi group. Taken together, these results suggest that RNAi might be an important tool for assessing the significance of tick salivary gland secreted proteins modulating responses at the tick-host interface.

IT INDEXING IN PROGRESS

CC 3 (Biochemical Genetics)

IT *Amblyomma americanum*

Protein sequences

cDNA sequences

(rNA interference in ticks, a study using **histamine binding protein** dsRNA in the female tick *Amblyomma americanum*)

IT 504653-46-7, GenBank AY246557

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
(Biological study)

(nucleotide sequence; rNA interference in ticks, a study using  
**histamine binding protein** dsRNA in the  
female tick *Amblyomma americanum*)

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS  
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 2 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:385176 HCAPLUS

DOCUMENT NUMBER: 137:229499

TITLE: Comparison of differentially expressed genes in the  
salivary glands of male ticks, *Amblyomma americanum*  
and *Dermacentor andersoni*

AUTHOR(S): Bior, Abdelaziz D.; Essenberg, Richard C.; Sauer, John  
R.

CORPORATE SOURCE: 246 Noble Research Center, Department of Biochemistry  
and Molecular Biology, Oklahoma State University,  
Stillwater, OK, 74078-3035, USA

**THIS PAGE BLANK (USPTO)**

SOURCE: Insect Biochemistry and Molecular Biology (2002),  
32(6), 645-655  
CODEN: IBMBES; ISSN: 0965-1748  
PUBLISHER: Elsevier Science Ltd.  
DOCUMENT TYPE: Journal  
LANGUAGE: English

- AB Genes expressed differentially in the salivary glands of unfed and fed male ticks, *Amblyomma americanum* (L.), were identified, cloned and sequenced, and some were compared with those expressed in the salivary glands of *Dermacentor andersoni*. Total protein and RNA increased sixfold in the salivary glands of fed male *A. americanum*, while in fed male *D. andersoni* salivary glands, RNA increased approx. 3.5 times. Feeding *D. andersoni* in the presence of females increased total RNA by 25% over those fed in the absence of females. Complementary DNAs were synthesized from RNA obtained from unfed and fed ticks and amplified using RNA arbitrarily primed polymerase chain reaction (RAP-PCR) with three different primers in sep. reactions. Differential display showed clear banding differences between the fed and the unfed ticks in *A. americanum* and *D. andersoni*. Sixty-one cDNA fragments that appeared to be from differentially expressed genes in *A. americanum* were isolated, cloned and sequenced. Hybridization reactions with labeled cDNA probes confirmed the differential expression of many of the genes in unfed and fed ticks' salivary glands; however, many of the bands contained more than one fragment and some of the fragments isolated from apparently differential bands were not specific. Sequences for 28 of the cDNA fragments (150-600 nucleotides in length) demonstrated similarity to genes in the databases, but nine of these were similar to sequences of unknown function. Some of the gene fragments identified may be important to tick feeding or tick salivary gland physiol., including a histamine-binding protein, an org. ion transporter, an apoptosis inhibitor, a cathepsin-B-like cysteine protease, proteins involved in gene regulation and several proteins involved in protein synthesis. Cross-hybridization of identified cDNAs from *A. americanum* with cDNA probes synthesized from *D. andersoni* total RNA did not show significant similarity between the two species.
- CC 12-2 (Nonmammalian Biochemistry)  
Section cross-reference(s): 3
- ST gene expression salivary gland tick *Amblyomma* *Dermacentor* male feeding;  
sequence expressed sequence tag male feeding salivary gland tick
- IT *Amblyomma americanum*  
*Dermacentor andersoni*  
Salivary gland  
Sex  
(comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT mRNA  
RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT EST (expressed sequence tag)  
Gene, animal  
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
(comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT cDNA sequences  
(for differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT Blood  
(meal of; comparison of differentially expressed genes in the salivary

**THIS PAGE BLANK (USPTO)**

glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)  
 IT Proteins  
 RL: BSU (Biological study, unclassified); BIOL (Biological study)  
 (salivary gland; comparison of differentially expressed genes in the  
 salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor  
 andersoni*)

IT	347782-65-4, GenBank BI273529	347782-66-5, GenBank BI273530
	347782-67-6, GenBank BI273531	347782-68-7, GenBank BI273532
	347782-69-8, GenBank BI273542	347782-70-1, GenBank BI273543
	347782-71-2, GenBank BI273544	347782-72-3, GenBank BI273545
	347782-73-4, GenBank BI273546	347782-74-5, GenBank BI273547
	347782-75-6, GenBank BI273548	347782-76-7, GenBank BI273549
	347782-77-8, GenBank BI273550	347782-78-9, GenBank BI273533
	347782-79-0, GenBank BI273534	347782-80-3, GenBank BI273535
	347782-81-4, GenBank BI273536	347782-82-5, GenBank BI273537
	347782-83-6, GenBank BI273538	347782-84-7, GenBank BI273539
	347782-85-8, GenBank BI273540	347782-86-9, GenBank BI273541
	347782-87-0, GenBank BI273551	347782-88-1, GenBank BI273552
	347782-89-2, GenBank BI273554	347782-90-5, GenBank BI273555
	347782-91-6, GenBank BI273556	347782-92-7, GenBank BI273557
	347782-93-8, GenBank BI273558	347782-94-9, GenBank BI273559
	347782-95-0, GenBank BI273560	347782-96-1, GenBank BI273561
	347782-97-2, GenBank BI273562	347782-98-3, GenBank BI273563
	347782-99-4, GenBank BI273564	347783-00-0, GenBank BI273565
	347783-01-1, GenBank BI273566	347783-02-2, GenBank BI273567
	347783-03-3, GenBank BI273568	347783-04-4, GenBank BI273569
	347783-05-5, GenBank BI273570	347783-06-6, GenBank BI273571
	347783-07-7, GenBank BI273572	347783-08-8, GenBank BI273573
	347783-09-9, GenBank BI273574	347783-10-2, GenBank BI273575
	347783-11-3, GenBank BI273576	347783-12-4, GenBank BI273577
	347783-13-5, GenBank BI273578	347783-14-6, GenBank BI273579
	347783-15-7, GenBank BI273580	347783-16-8, GenBank BI273581
	347783-17-9, GenBank BI273582	347783-18-0, GenBank BI273583
	347783-19-1, GenBank BI273584	347783-20-4, GenBank BI273585
	347783-21-5, GenBank BI273586	347783-22-6, GenBank BI273587
	347783-23-7, GenBank BI273588	347783-24-8, GenBank BI273589
	347783-25-9, GenBank BI273590	347783-26-0, GenBank BI273591
	347783-27-1, GenBank BI273592	347783-28-2, GenBank BI273593
	347783-29-3, GenBank BI273594	347783-30-6, GenBank BI273595
	347783-31-7, GenBank BI273596	347783-32-8, GenBank BI273597
	347783-33-9, GenBank BI273598	347783-34-0, GenBank BI273599
	347783-35-1, GenBank BI273600	347783-36-2, GenBank BI273601
	347783-37-3, GenBank BI273602	347783-38-4, GenBank BI273603
	347783-39-5, GenBank BI273604	347783-40-8, GenBank BI273605
	347783-41-9, GenBank BI273606	347783-42-0, GenBank BI273607
	347783-43-1, GenBank BI273608	347783-44-2, GenBank BI273609
	347783-45-3, GenBank BI273610	347783-46-4, GenBank BI273611
	347783-47-5, GenBank BI273612	347783-48-6, GenBank BI273613
	347783-49-7, GenBank BI273614	348411-30-3, GenBank BI275494
	348411-31-4, GenBank BI275495	348411-32-5, GenBank BI275496
	348411-33-6, GenBank BI275497	348411-34-7, GenBank BI275498
	348411-35-8, GenBank BI275499	348411-36-9, GenBank BI275500
	348411-37-0, GenBank BI275501	348411-38-1, GenBank BI275502
	348411-39-2, GenBank BI275503	348411-40-5, GenBank BI275504
	348411-41-6, GenBank BI275505	348411-42-7, GenBank BI275506
	348411-43-8, GenBank BI275507	348411-44-9, GenBank BI275508
	348411-45-0, GenBank BI275509	348411-46-1, GenBank BI275510
	348411-47-2, GenBank BI275511	348411-48-3, GenBank BI275512
	348411-49-4, GenBank BI275513	386326-31-4, GenBank BI273553

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

**THIS PAGE BLANK (USPTO)**

## (Biological study)

(nucleotide sequence; comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)

REFERENCE COUNT: 61 THERE ARE 61 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 3 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:231098 HCAPLUS

DOCUMENT NUMBER: 136:383078

TITLE: A high affinity serotonin- and histamine-binding lipocalin from tick saliva

AUTHOR(S): Sangamnatdej, S.; Paesen, G. C.; Slovak, M.; Nuttall, P. A.

CORPORATE SOURCE: CEH Oxford, Oxford, OX1 3SR, UK

SOURCE: Insect Molecular Biology (2002), 11(1), 79-86

CODEN: IMBIE3; ISSN: 0962-1075

PUBLISHER: Blackwell Publishing Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB To overcome the inflammatory response in its host, the cattle-feeding, brown ear tick secretes **histamine-binding proteins** into the feeding site. These proteins are .beta.-barrels with 2 internal binding sites: a high-affinity (H) site for histamine and a site (L) for which the natural ligand is unknown. Here we report a related protein (SHBP), secreted by a rodent- and cattle-feeding tick, that traps both histamine and serotonin. The histamine-binding H site is well conserved in SHBP, whereas residue changes in the L-like site are consistent with binding of the bulkier serotonin mol. As histamine is a key inflammatory mediator in cattle, while serotonin takes on this role in rodents, the diversification of these tick proteins may reflect host adaptation.

CC 12-1 (Nonmammalian Biochemistry)

Section cross-reference(s): 3, 6

ST Dermacentor saliva serotonin histamine binding lipocalin sequence

IT Dermacentor reticulatus

Feeding

Protein sequences

Saliva

cDNA sequences

(high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT Proteins

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)

(lipocalin; high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT 426310-26-1, Lipocalin (Dermacentor reticulatus)

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)

(amino acid sequence; high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT 50-67-9, Serotonin, biological studies 51-45-6, Histamine, biological studies

RL: BSU (Biological study, unclassified); BIOL (Biological study)

(high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT 381905-30-2, GenBank AF217101

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)

**THIS PAGE BLANK (USPTO)**



These are all of the cites for the Registry #5 for FSHBP1, 2, MSHBP1 & D.RET6

FSHBP1

SRIVASTAVA 10/087,195

FSHBP1

FSHBP-2

=> d que 130

L5

4 SEA FILE=REGISTRY ABB=ON PLU=ON (329085-08-7/BI OR 329085-09-8/BI OR ~~329085-10-1/BI OR 51-45-6/BI~~)

L6

3 SEA FILE=REGISTRY ABB=ON PLU=ON L5 AND PROTEIN/FS

L9

3 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTACALIN? 3 cites for Histacalin

L10

2 SEA FILE=HCAPLUS ABB=ON PLU=ON L6 2 cites for FSHBP1, 2 & MSHBP1 from Registry

L11

4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP1

L13

4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP2

L14

4 SEA FILE=HCAPLUS ABB=ON PLU=ON MS-HBP1

L15

4 SEA FILE=HCAPLUS ABB=ON PLU=ON "D.RET6"

L18

4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-1

L19

4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-2

L24

6 SEA FILE=HCAPLUS ABB=ON PLU=ON (L9 OR L10) OR L11 OR (L13 OR L14 OR L15) OR (L18 OR L19) 6 cites total

L25

31 SEA FILE=REGISTRY ABB=ON PLU=ON (51-45-6/BI OR 200220-28-6/BI OR 200220-29-7/BI OR 200220-30-0/BI OR 200220-32-2/BI OR 200220-33-3/BI OR 200220-34-4/BI OR 200220-31-1/BI OR 200220-35-5/BI OR 329085-08-7/BI OR 329085-10-1/BI OR 226219-91-6/BI OR 226890-87-5/BI OR 226890-96-6/BI OR 226890-97-7/BI OR 226890-98-8/BI OR 226891-00-5/BI OR 226891-16-3/BI OR 226891-31-2/BI OR 226891-34-5/BI OR 226891-36-7/BI OR 226891-37-8/BI OR 226891-38-9/BI OR 226891-39-0/BI OR 226891-40-3/BI OR 226927-07-7/BI OR 329085-09-8/BI OR 329085-11-2/BI OR 50-67-9/BI OR 501-75-7/BI OR 644-42-8/BI)

31 compounds in the combined L24 cites

L26

16 SEA FILE=REGISTRY ABB=ON PLU=ON L25 AND PROTEIN/FS 16 are proteins

L27

5 SEA FILE=HCAPLUS ABB=ON PLU=ON L24 AND L26 5 cites

L28

1 SEA FILE=REGISTRY ABB=ON PLU=ON L26 AND SQL=209 ← D.RET6

L29

2 SEA FILE=HCAPLUS ABB=ON PLU=ON L28 2 cites for

L30

5 SEA FILE=HCAPLUS ABB=ON PLU=ON L29 OR L27 5 cites total

FS = file segment

**THIS PAGE BLANK (USPTO)**

=&gt; d ibib abs hitstr ind 1-5 130

L30 ANSWER 1 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:168020 HCAPLUS

DOCUMENT NUMBER: 134:217189

TITLE: Treatment of allergic rhinitis with proteins from ticks

INVENTOR(S): Nuttall, Patricia Anne; Paesen, Guido Christiaan

PATENT ASSIGNEE(S): Evlutec Limited, UK

SOURCE: PCT Int. Appl., 19 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001016164	A2	20010308	WO 2000-GB3287	20000824
WO 2001016164	A3	20010503		

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

BR 2000013655	A	20020507	BR 2000-13655	20000824
---------------	---	----------	---------------	----------

EP 1207899	A2	20020529	EP 2000-954788	20000824
------------	----	----------	----------------	----------

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL

JP 2003508410	T2	20030304	JP 2001-519725	20000824
---------------	----	----------	----------------	----------

US 2002193306	A1	20021219	US 2002-87195	20020301
---------------	----	----------	---------------	----------

PRIORITY APPLN. INFO.:

GB 1999-20673 A 19990901

WO 2000-GB3287 W 20000824

AB The invention relates to the discovery that various proteins isolated from ticks are effective in the treatment and prevention of allergic rhinitis. These proteins may most suitably be applied to an effected area and are thus effective to treat this condition and to ameliorate its symptoms. Human subjects were challenged with histamine and then were treated with histamine-binding protein, MS-HBP1.

IT 329085-08-7

RL: PRP (Properties)

(unclaimed protein sequence; treatment of allergic rhinitis with proteins from ticks)

RN 329085-08-7 HCAPLUS

CN 1: PN: WO0116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IT 329085-09-8 329085-10-1

RL: PRP (Properties)

(unclaimed sequence; treatment of allergic rhinitis with proteins from ticks)

RN 329085-09-8 HCAPLUS

CN 2: PN: WO0116164 TABLE: 1 unclaimed sequence (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

**THIS PAGE BLANK (USPTO)**

RN 329085-10-1 HCAPLUS

CN 3: PN: W00116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IC ICM C07K014-00

CC 1-7 (Pharmacology)

Section cross-reference(s): 12, 15

ST allergic rhinitis treatment protein tick; histamine binding protein MSHBP1  
treatment allergic rhinitis

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(D.RET6; treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(FS-HBP1 (female-specific histamine-binding protein 1); treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(FS-HBP2 (female-specific histamine-binding protein 2); treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(MS-HBP1 (male-specific histamine-binding protein 1); treatment of allergic rhinitis with proteins from ticks)

IT Nose

(allergic rhinitis; treatment of allergic rhinitis with proteins from ticks)

IT Nervous system stimulants

(antisedatives, medicament also contg.; treatment of allergic rhinitis with proteins from ticks)

IT Parasite

(ecto-, histacal in of; treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(histacalins; treatment of allergic rhinitis with proteins from ticks)

IT Antihistamines

(medicament also contg.; treatment of allergic rhinitis with proteins from ticks)

IT Proteins, general, biological studies

RL: THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(of ticks; treatment of allergic rhinitis with proteins from ticks)

IT Allergy inhibitors

Drug delivery systems

Hay fever

**THIS PAGE BLANK (USPTO)**

## Mite and Tick

(treatment of allergic rhinitis with proteins from ticks)

IT 51-45-6, Histamine, biological studies

RL: ADV (Adverse effect, including toxicity); BPR (Biological process);  
 BSU (Biological study, unclassified); BIOL (Biological study); PROC  
 (Process)

(proteins binding to; treatment of allergic rhinitis with proteins from ticks)

IT 329085-08-7

RL: PRP (Properties)

(unclaimed protein sequence; treatment of allergic rhinitis with proteins from ticks)

IT 329085-09-8 329085-10-1

RL: PRP (Properties)

(unclaimed sequence; treatment of allergic rhinitis with proteins from ticks)

L30 ANSWER 2 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:167826 HCAPLUS

DOCUMENT NUMBER: 134:217188

TITLE: use of histacalin protein for treatment or prevention of conjunctivitis

INVENTOR(S): Nuttall, Patricia Anne; Paesen, Guido Christiaan

PATENT ASSIGNEE(S): Evolutec Limited, UK

SOURCE: PCT Int. Appl., 19 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

## PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001015719	A2	20010308	WO 2000-GB3282	20000824
WO 2001015719	A3	20010510		

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

BR 2000013665 A 20020514 BR 2000-13665 20000824

EP 1207898 A2 20020529 EP 2000-954784 20000824

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL

JP 2003508445 T2 20030304 JP 2001-519931 20000824

US 2002151499 A1 20021017 US 2002-85572 20020227

PRIORITY APPLN. INFO.: GB 1999-20674 A 19990901

WO 2000-GB3282 W 20000824

AB Various histacalin proteins isolated from ticks are effective in the treatment of conjunctivitis. These proteins may most suitably be applied topically to an affected area and are effective to ameliorate the symptoms of this condition.

IT 329085-08-7 329085-10-1 329085-11-2

RL: PRP (Properties)

(unclaimed sequence; use of histacalin protein for treatment or prevention of conjunctivitis)

**THIS PAGE BLANK (USPTO)**



RN 329085-08-7 HCAPLUS  
 CN 1: PN: W00116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 329085-10-1 HCAPLUS  
 CN 3: PN: W00116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 329085-11-2 HCAPLUS  
 CN 2: PN: W00115719 TABLE: 1 unclaimed sequence (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IC ICM A61K038-00  
 CC 1-7 (Pharmacology)  
 Section cross-reference(s): 63  
 ST **histacalin** antihistamine antiallergic antiinflammatory conjunctivitis  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
     (D.RET6; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
     (FS-HBP1; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
     (FS-HBP2; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
     (MS-HBP1; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Eye, disease  
     (allergic conjunctivitis; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Eye, disease  
     (conjunctivitis; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Mast cell  
     (degranulation; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Parasite  
     (ecto-, blood-feeding; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Allergy inhibitors  
     Anti-inflammatory agents  
     Antihistamines  
     Drug delivery systems  
     Mite and Tick  
     Protein sequences  
     (**histacalin** protein for treatment or prevention)

**THIS PAGE BLANK (USPTO)**

conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (histacalin; histacalin protein for treatment or prevention conjunctivitis)  
 IT Cell degranulation  
 (mast cell; histacalin protein for treatment or prevention conjunctivitis)  
 IT Mental activity  
 (sedation, antisedatives; histacalin protein for treatment or prevention conjunctivitis)  
 IT Drug delivery systems  
 (solns., ophthalmic; histacalin protein for treatment or prevention conjunctivitis)  
 IT 329085-08-7 329085-10-1 329085-11-2  
 RL: PRP (Properties)  
 (unclaimed sequence; use of histacalin protein for treatment or prevention of conjunctivitis)

L30 ANSWER 3 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:374997 HCAPLUS

DOCUMENT NUMBER: 131:154999

TITLE: Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure

AUTHOR(S): Paesen, G. C.; Adams, P. L.; Harlos, K.; Nuttall, P. A.; Stuart, D. I.

CORPORATE SOURCE: Institute of Virology and Environmental Microbiology, Natural Environment Research Council, Oxford, OX1 3SR, UK

SOURCE: Molecular Cell (1999), 3(5), 661-671

CODEN: MOCEFL; ISSN: 1097-2765

PUBLISHER: Cell Press

DOCUMENT TYPE: Journal

LANGUAGE: English

AB High-affinity histamine-binding proteins (HBPs) were discovered in the saliva of Rhipicephalus appendiculatus ticks. Their ability to outcompete histamine receptors indicates that they suppress inflammation during blood feeding. The crystal structure of a histamine-bound HBP, detd. at 1.25 Å. resoln., reveals a lipocalin fold novel in contg. two binding sites for the same ligand. The sites are orthogonally arranged and highly rigid and form an internal surface of unusual polar character that complements the physicochem. properties of histamine. As sol. receptors of histamine, HBPs offer a new strategy for controlling histamine-based diseases.

IT 200220-32-2 200220-33-3 200220-34-4

RL: PRP (Properties)

(amino acid sequence; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

RN 200220-32-2 HCAPLUS

CN Protein FS-HBP1 (Rhipicephalus appendiculatus female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-33-3 HCAPLUS

CN Protein FS-HBP2 (Rhipicephalus appendiculatus female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

**THIS PAGE BLANK (USPTO)**

RN 200220-34-4 HCAPLUS

CN Protein MS-HBP1 (Rhipicephalus appendiculatus male-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

CC 6-3 (General Biochemistry)

Section cross-reference(s): 3, 12

ST Rhipicephalus histamine binding protein structure; cDNA sequence tick histamine binding protein

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(HBP (histamine-binding proteins); isolation, cloning, mol.

characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(HBP-1 (histamine-binding protein

1); isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(HBP-2 (histamine-binding protein

2); isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Crystal structure

Protein sequences

Rhipicephalus appendiculatus

cDNA sequences

(isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Tertiary structure

(protein; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT 200220-32-2 200220-33-3 200220-34-4

RL: PRP (Properties)

(amino acid sequence; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT 51-45-6, Histamine, biological studies

RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)

(isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT 200220-28-6, GenBank U96080 200220-29-7, GenBank U96081 200220-30-0, GenBank U96082

RL: PRP (Properties)

(nucleotide sequence; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

REFERENCE COUNT: 48 THERE ARE 48 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

**THIS PAGE BLANK (USPTO)**

L30 ANSWER 4 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:359659 HCAPLUS  
 DOCUMENT NUMBER: 131:28315  
 TITLE: Cloning and functions of vasoactive amine-binding proteins from ticks  
 INVENTOR(S): Nuttall, Patricia Ann; Paesen, Guido Christian  
 PATENT ASSIGNEE(S): Oxford Vacs Ltd., UK  
 SOURCE: PCT Int. Appl., 84 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9927104	A1	19990603	WO 1998-GB3530	19981126
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, VZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
CA 2309809	AA	19990603	CA 1998-2309809	19981126
AU 9912511	A1	19990615	AU 1999-12511	19981126
EP 1034273	A1	20000913	EP 1998-955786	19981126
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI, RO				
BR 9815056	A	20001003	BR 1998-15056	19981126
JP 2002508927	T2	20020326	JP 2000-522246	19981126
NZ 504753	A	20021126	NZ 1998-504753	19981126
PRIORITY APPLN. INFO.:			GB 1997-25046	A 19971126
			GB 1998-13917	A 19980626
			WO 1998-GB3530	W 19981126

AB The present invention relates to histamine and serotonin binding mols. that possess a binding site with the precise mol. configuration that is necessary to confer on the mol. a high affinity for histamine. The invention includes proteins, peptides and chem. compds. that possess this mol. configuration and that are thus able to bind to histamine with high affinity. These mols. may be used in the regulation of the action of histamine or serotonin, the detection and quantification of histamine or serotonin and in the treatment of various diseases and allergies. The mols. may also be used as components of vaccines directed against blood-sucking ectoparasites. Vasoactive amine binding proteins (VABPs) are provided that specifically bind to vasoactive amines with a dissocn. const. of  $<10^{-7}$  M and which belong to the same protein family as MS-HBP1, FS-HBP1, FS-HBP2 and D.RET6. Thus, 11 VASPs were isolated, and their cDNAs cloned and sequenced, from ticks: FS-HBP1 (female-specific histamine-binding protein 1), FS-HBP2 (female-specific histamine-binding protein 2), MS-HBP1 (male-specific histamine-binding protein 1), and Ra-Res from Rhipicephalus appendiculatus; D.RET6 from Dermacentor reticularis; Av-HBP from Amblyomma variegatum; and 5 related Ih/Bm-HBP proteins from a mixed Ixodes hexagonus/Boophilus microplus cDNA expression

**THIS PAGE BLANK (USPTO)**



library. These VASPs possess similar amino acid sequences and predicted secondary structures. The VASPs bind histamine in mammals, and can be used as anti-inflammatory agents to regulate histamine action and to control its pathol. effects. The crystal structure of FS-HBP2 to 2.24 Å. resoln. was used to design a synthetic cyclic octapeptide (-Ala-Glu-Ala-Phe-Ala-Glu-Ala-Trp-) with histamine binding activity.

IT 200220-32-2 200220-33-3 200220-34-4  
200220-35-5 226891-31-2 226891-34-5  
226891-36-7 226891-37-8 226891-38-9  
226891-39-0 226891-40-3

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

RN 200220-32-2 HCAPLUS

CN Protein FS-HBP1 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-33-3 HCAPLUS

CN Protein FS-HBP2 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-34-4 HCAPLUS

CN Protein MS-HBP1 (*Rhipicephalus appendiculatus* male-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-35-5 HCAPLUS

CN Protein D.RET6 (*Dermacentor reticularis* histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-31-2 HCAPLUS

CN Histamine-binding protein Ra-Res (*Rhipicephalus appendiculatus*) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-34-5 HCAPLUS

CN Histamine-binding protein Av-HBP (*Amblyomma variegatum*) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-36-7 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP1 (*Boophilus microplus*/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-37-8 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP2 (*Boophilus microplus*/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-38-9 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP3 (*Boophilus microplus*/Ixodes hexagonus) (9CI) (CA INDEX NAME)

43 PAGE BLANK (USPTO)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-39-0 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP4 (Boophilus microplus/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-40-3 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP5 (Boophilus microplus/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IT 226219-91-6

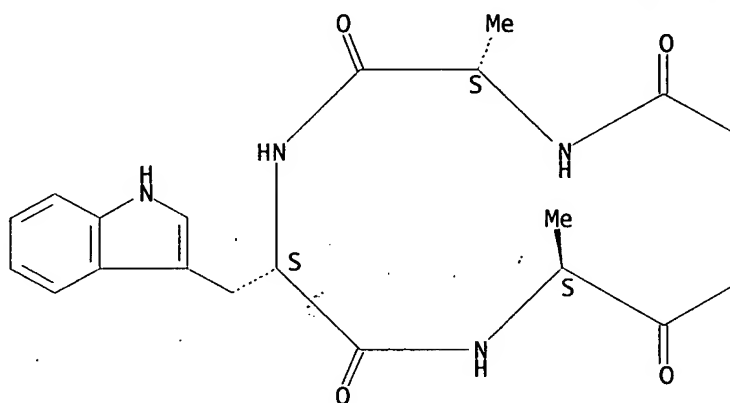
RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (cloning and functions of vasoactive amine-binding proteins from ticks)

RN 226219-91-6 HCAPLUS

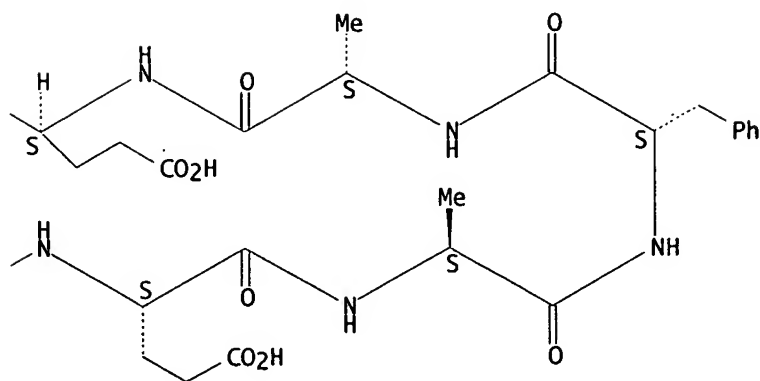
CN Cyclo(L-alanyl-L-.alpha.-glutamyl-L-alanyl-L-phenylalanyl-L-alanyl-L-.alpha.-glutamyl-L-alanyl-L-tryptophyl) (9CI) (CA INDEX NAME)

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



**THIS PAGE BLANK (USPTO)**

- IC ICM C12N015-21  
ICS C07K014-435; A61K038-17; A61K031-40; A61K031-19; A61K031-66;  
A61K031-35; G01N033-68; A23L001-015; C12N005-10; C12N001-21;  
A01K067-027; C12N015-00
- CC 2-8 (Mammalian Hormones)  
Section cross-reference(s): 1, 3, 6, 12, 75
- ST vasoactive amine binding protein tick; histamine binding protein tick;  
Rhipicephalus vasoactive amine binding protein; Dermacenter vasoactive  
amine binding protein; Amblyomma vasoactive amine binding protein;  
Boophilus vasoactive amine binding protein; Ixodes vasoactive amine  
binding protein; sequence vasoactive binding protein cDNA tick;  
inflammation inhibitor VASP protein tick
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(D.RET6; cloning and functions of vasoactive  
amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(FS-HBP1 (female-specific histamine-  
binding protein 1); cloning and functions  
of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(FS-HBP2 (female-specific histamine-  
binding protein 2); cloning and functions  
of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(Ih/Bm-HBP (Ixodes hexagonus/Boophilus microplus histamine-binding  
protein); cloning and functions of vasoactive amine-binding proteins  
from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(MS-HBP1 (male-specific histamine-  
binding protein 1); cloning and functions  
of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(VASP (vasoactive amine-binding protein); cloning and functions of  
vasoactive amine-binding proteins from ticks)
- IT Amines, biological studies  
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL  
(Biological study); PROC (Process)  
(biogenic, vasoactive; cloning and functions of vasoactive  
amine-binding proteins from ticks)
- IT 5-HT antagonists  
Amblyomma variegatum

**THIS PAGE BLANK (USPTO)**

- Anti-inflammatory agents
- Antihistamines
- Boophilus microplus
- Dermacentor reticulatus
- Ixodes hexagonus
- Mite and Tick
- Molecular cloning
- Rhipicephalus appendiculatus
- Scorpion
- Snake
- Spider
- Transformation, genetic
- Vaccines
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Fusion proteins (chimeric proteins)
  - RL: BPN (Biosynthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Peptides, biological studies
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (cyclic; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Parasite
  - (ecto-; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT cDNA sequences
  - (for vasoactive amine-binding proteins from ticks)
- IT Food analysis
- Plant analysis
  - (histamine detn. in; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Diagnosis
  - (mol.; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Crystal structure
  - (of histamine-binding protein FS-HBP2; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Protein sequences
  - (of vasoactive amine-binding proteins from ticks)
- IT Conformation
  - (protein, of histamine-binding protein FS-HBP2; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Body fluid
  - (quantification of histamine levels in body fluids; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 200220-32-2 200220-33-3 200220-34-4  
 200220-35-5 226891-31-2 226891-34-5  
 226891-36-7 226891-37-8 226891-38-9  
 226891-39-0 226891-40-3
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 50-67-9, Serotonin, biological studies 51-45-6, Histamine, biological studies 501-75-7, 1-Methylhistamine 644-42-8, 3-Methylhistamine
  - RL: ANT (Analyte); BPR (Biological process); BSU (Biological study, unclassified); ANST (Analytical study); BIOL (Biological study); PROC

**THIS PAGE BLANK (USPTO)**



(Process)

(cloning and functions of vasoactive amine-binding proteins from ticks)

IT 226219-91-6

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)

(cloning and functions of vasoactive amine-binding proteins from ticks)

IT 200220-28-6P 200220-29-7P 200220-30-0P 200220-31-1P 226890-87-5P  
226890-96-6P 226890-97-7P 226890-98-8P 226891-00-5P 226891-16-3P  
226927-07-7P

RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(nucleotide sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L30 ANSWER 5 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1997:776255 HCAPLUS

DOCUMENT NUMBER: 128:57765

TITLE: Cloning and functions of vasoactive amine-binding proteins from ticks

INVENTOR(S): Paesen, Guido Christian; Nuttall, Patricia Ann

PATENT ASSIGNEE(S): Oxford Vacs Ltd., UK; Paesen, Guido Christian; Nuttall, Patricia Ann

SOURCE: PCT Int. Appl., 44 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9744451	A2	19971127	WO 1997-GB1372	19970519
WO 9744451	A3	19980219		

W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

CA 2253924	AA	19971127	CA 1997-2253924	19970519
------------	----	----------	-----------------	----------

AU 9729071	A1	19971209	AU 1997-29071	19970519
------------	----	----------	---------------	----------

AU 725630	B2	20001019		
-----------	----	----------	--	--

EP 906425	A2	19990407	EP 1997-923204	19970519
-----------	----	----------	----------------	----------

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI, RO

BR 9709101	A	19990803	BR 1997-9101	19970519
------------	---	----------	--------------	----------

CN 1225683	A	19990811	CN 1997-196317	19970519
------------	---	----------	----------------	----------

NZ 332648	A	20000526	NZ 1997-332648	19970519
-----------	---	----------	----------------	----------

JP 2000512489	T2	20000926	JP 1997-541799	19970519
---------------	----	----------	----------------	----------

PRIORITY APPLN. INFO.:			GB 1996-10484	A	19960518
------------------------	--	--	---------------	---	----------

			GB 1997-7844	A	19970418
--	--	--	--------------	---	----------

			WO 1997-GB1372	W	19970519
--	--	--	----------------	---	----------

AB Vasoactive amine binding proteins (VABPs) are provided that specifically bind to vasoactive amines with a dissociation constant of  $<10^{-7}$  M and which belong to the same protein family as MS-HBP1,

**THIS PAGE BLANK (USPTO)**

FS-HBP1, FS-HBP2 and D.

RET6. Thus, 4 VASPs were isolated, and their cDNAs cloned and sequenced, from ticks: FS-HBP1 (female-specific histamine-binding protein 1), FS-HBP2 (female-specific histamine-binding protein 2), and MS-HBP1 (male-specific histamine-binding protein 1) from *Rhipicephalus appendiculatus*; and D.RET6 from *Dermacenter reticularis*. These 4 VASPs possess similar amino acid sequences and predicted secondary structures. The VASPs bind histamine in mammals, and can be used as anti-inflammatory agents to regulate histamine action and to control its pathol. effects.

IT 200220-32-2 200220-33-3 200220-34-4  
200220-35-5

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

RN 200220-32-2 HCAPLUS

CN Protein FS-HBP1 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-33-3 HCAPLUS

CN Protein FS-HBP2 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-34-4 HCAPLUS

CN Protein MS-HBP1 (*Rhipicephalus appendiculatus* male-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-35-5 HCAPLUS

CN Protein D.RET6 (*Dermacenter reticularis* histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IC ICM C12N015-12

ICS C12N015-86; A01K067-027; A61K038-17; C12N005-10; G01N033-68; C12N015-62; C07K014-435

CC 2-8 (Mammalian Hormones)

Section cross-reference(s): 1, 3

ST vasoactive amine binding protein tick; histamine binding protein tick; *Rhipicephalus* vasoactive amine binding protein; *Dermacenter* vasoactive amine binding protein; sequence vasoactive binding protein cDNA tick; inflammation inhibitor VASP protein tick

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (D.RET6; cloning and functions of vasoactive amine-binding proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (FS-HBP1 (female-specific histamine-binding protein 1); cloning and functions

**THIS PAGE BLANK (USPTO)**

- of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (FS-HBP2 (female-specific histamine-binding protein 2); cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (MS-HBP1 (male-specific histamine-binding protein 1); cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (VASP (vasoactive amine-binding protein); cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Amines, biological studies
  - RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
  - (biogenic, vasoactive; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Anti-inflammatory agents
  - Antihistamines
  - Dermacentor reticulatus
  - Mite and Tick
  - Molecular cloning
  - Rhipicephalus appendiculatus
  - Scorpion
  - Snake
  - Spider
  - Transformation, genetic
  - Vaccines
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Fusion proteins (chimeric proteins)
  - RL: BPN (Biosynthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Parasite
  - (ecto-; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT cDNA sequences
  - (for vasoactive amine-binding proteins from ticks)
- IT Protein sequences
  - (of vasoactive amine-binding proteins from ticks)
- IT Body fluid
  - (quantification of histamine levels in body fluids; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 200220-32-2 200220-33-3 200220-34-4 200220-35-5
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 51-45-6, Histamine, biological studies

**THIS PAGE BLANK (USPTO)**

RL: ANT (Analyte); BPR (Biological process); BSU (Biological study, unclassified); ANST (Analytical study); BIOL (Biological study); PROC (Process)

(cloning and functions of vasoactive amine-binding proteins from ticks)

IT 200220-28-6P 200220-29-7P 200220-30-0P 200220-31-1P

RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(nucleotide sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

**THIS PAGE BLANK (USPTO)**



(nucleotide sequence; high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

REFERENCE COUNT: 25 THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 4 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:8145 HCAPLUS

DOCUMENT NUMBER: 134:128741

TITLE: Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods

AUTHOR(S): Montfort, William R.; Weichsel, Andrzej; Andersen, John F.

CORPORATE SOURCE: Department of Biochemistry, University of Arizona, Tucson, AZ, 85721, USA

SOURCE: Biochimica et Biophysica Acta (2000), 1482(1-2), 110-118

CODEN: BBACAQ; ISSN: 0006-3002

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal; General Review

LANGUAGE: English

AB A review, with 53 refs. Recent gene sequence and crystal structure detns. of salivary proteins from several blood-sucking arthropods have revealed an unusual evolutionary relationship: many such proteins derive their functions from lipocalin protein folds. Many blood-sucking arthropods have independently evolved the ability to overcome a host organism's means of preventing blood loss (called hemostasis). Most blood feeders have proteins that induce vasodilation, inhibit blood coagulation, and reduce inflammation, but do so by distinctly different mechanisms. Despite this diversity, in many cases the antihemostatic activities in such organisms reside in proteins with lipocalin folds. Thirteen such lipocalins are described in this review, with a particular focus on the heme-contg. nitrophorins from *R. prolixus*, which transport nitric oxide, sequester histamine, and disrupt blood coagulation. Also described are the anti-platelet compds. RPAI, moubatin, and pallidipin from *R. prolixus*, *Ornithodoros moubata*, and *Triatoma pallidipennis*; the antithrombin protein triabin from *T. pallidipennis*; and the tick histamine binding proteins from *Rhipicephalus appendiculatus*.

CC 12-0 (Nonmammalian Biochemistry)

Section cross-reference(s): 6

ST review blood sucking arthropod nitrophorin lipocalin

IT Arthropod (Arthropoda)

(blood-sucking; nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(lipocalins; nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

IT *Rhodnius prolixus*

Salivary gland

(nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(nitrophorins; nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

**THIS PAGE BLANK (USPTO)**

REFERENCE COUNT: 53 THERE ARE 53 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 5 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:8143 HCAPLUS

DOCUMENT NUMBER: 134:128740

TITLE: Tick **histamine-binding**

**proteins:** lipocalins with a second binding cavity

AUTHOR(S): Paesen, Guido C.; Adams, Peter L.; Nuttall, Patricia A.; Stuart, David L.

CORPORATE SOURCE: CEH Institute of Virology and Environmental Microbiology, Oxford, OX1 3SR, UK

SOURCE: Biochimica et Biophysica Acta (2000), 1482(1-2), 92-101

CODEN: BBACAQ; ISSN: 0006-3002

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal; General Review

LANGUAGE: English

AB A review, with 26 refs. Tick **histamine-binding**

**proteins** (HBPs) are lipocalins with 2 binding pockets. One of these binds histamine with a high affinity and is found at the position expected from other lipocalins, adjacent to the .OMEGA.-loop at the open-end of the .beta.-barrel. A 2nd binding cavity, which is a low-affinity site for histamine in 1 of the HBPs, is located at the end of the barrel that is closed off in other lipocalins. To create the 2nd site, the closed-end region has undergone a major reconstruction. Typical lipocalin characteristics, such as the 310 helix and a structural cluster of highly conserved residues, have been lost, while an .alpha.-helix now shields the cavity from the exterior. The prominence of acidic residues in the binding pockets is another distinctive characteristic of HBPs. Whereas most lipocalins have highly hydrophobic binding cavities designed to bind lipophilic compds., HBPs have evolved to trap cationic, hydrophilic mols.

CC 12-0 (Nonmammalian Biochemistry)

Section cross-reference(s): 6

ST review tick **histamine binding protein**

IT Proteins, specific or class

RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)  
(histamine-binding; tick **histamine-binding proteins**)

IT Mite and Tick

(tick **histamine-binding proteins**)

REFERENCE COUNT: 26 THERE ARE 26 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 6 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1994:236350 HCAPLUS

DOCUMENT NUMBER: 120:236350

TITLE: Purification of histamine receptor proteins from detergent-solubilized human peripheral blood mononuclear cells

AUTHOR(S): Warlow, Robert S.; Rajasekariah, Poornima; Lambert, Paul; Morgan, Janelle; Dao, Lan Phuong; Bernard, Claude C. A.; Walls, Ronald S.

CORPORATE SOURCE: Immunology Department, Concord Repatriation Hospital, Sydney, Australia

SOURCE: Biochemistry (1994), 33(16), 4800-11  
CODEN: BICHAW; ISSN: 0006-2960

**THIS PAGE BLANK (USPTO)**

DOCUMENT TYPE: Journal  
 LANGUAGE: English

AB Histamine is released from mast cells and basophils by either immunol. or nonimmunol. mechanisms. Histamine, which is the most potent short acting mediator released from these cells, exerts its diverse biol. actions by binding to cell surface histamine receptors. The authors report the affinity purifn. of histamine receptor proteins from Triton X-100 solubilized peripheral human blood mononuclear cells which include lymphocytes and monocytes. Three different designs of histamine affinity columns were constructed; all three resulted in the same material being eluted. This consisted of bands which on SDS-PAGE after boiling and redn. had the following mol. wts.: 193K, 84K, 58K, 48K, 37K, and 16K. The most abundant bands were of mol. wts. 193K, 48K, and 16K, and these were disulfide bonded together to form a high mol. wt. complex. The 58K band was present in lower amts. than the others, and in only a few fractions. It had the same mol. wt. as the dimeric form of histamine methyltransferase which is present in small amts. in mononuclear cells and may therefore have copurified. The **histamine binding proteins** described in this report were purified by conventional affinity chromatog., rather than by an expression cloning approach which obviates the use of any protein chem. Consequently, the authors had the advantage of being able to verify the histamine binding specificity of the purified proteins directly and with several independent assays as follows. The histamine binding specificity of all three columns was established by specific elution with histamine, by preabsorption of crude cell ext. with excess free histamine prior to column application, and by comparison with control columns. Independent detn. of the binding specificity, using a radioreceptor dot blot assay, of the eluate contg. only the 193K, 48K, and 16K disulfide-linked subunits confirmed that the purified material bound specifically to [3H]histamine and that a 300-500-fold degree of purifn. from tissue ext. had been obtained. Following cell surface radioreceptor crosslinking of radiolabeled histamine to intact mononuclear cells, the 16K band was detected, indicating it to be the ligand-binding subunit for histamine. These same three proteins were purified from T lymphocyte and monocytoïd cell lines, indicating that both lymphocyte and monocyte subsets of mononuclear cells express these proteins. The trimol. structure, consisting of 193K, 48K, and 16K subunits appears to be a novel histamine cell surface receptor protein complex as the mol. wts. of the 193K, 48K, and 16K bands bore no relation to the predicted mol. wts. of the recently expression cloned H1 receptor gene derived from bovine adrenal and H2 receptor gene derived from human parietal cells; this may reflect receptor heterogeneity within a tissue, between tissues, and/or between species.

CC 2-2 (Mammalian Hormones)

ST histamine receptor mononuclear cell purifn characterization; affinity chromatog histamine receptor mononuclear cell; lymphocyte histamine receptor purifn characterization; monocyte histamine receptor purifn characterization; structure histamine receptor mononuclear cell

IT Monocyte

(histamine receptors of human, purifn. and characterization of)

IT Cell membrane

(histamine receptors of, of human blood mononuclear cells, purifn. and characterization of)

IT Conformation and Conformers

Disulfide group

(of histamine receptors, of human blood mononuclear cells)

IT Lymphocyte

(T-cell, histamine receptors of human, purifn. and characterization of)

IT Chromatography, column and liquid

(affinity, of histamine receptors, of human blood mononuclear cells)

**THIS PAGE BLANK (USPTO)**

- IT Receptors  
RL: BIOL (Biological study)  
(histaminic, of blood mononuclear cells, of human, purifn. and characterization of)
- IT Leukocyte  
(mononuclear, histamine receptors of human, purifn. and characterization of)
- IT 51-45-6P, Histamine, biological studies  
RL: BIOL (Biological study); PREP (Preparation)  
(receptor binding of, after purifn. from human blood mononuclear cells)

L32 ANSWER 7 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1991:421847 HCAPLUS  
DOCUMENT NUMBER: 115:21847  
TITLE: Histaglobulin and changes in conjunctival structure in vernal conjunctivitis  
AUTHOR(S): Volokhovskaya, Z. P.; Kadyrova, M. Ya.  
CORPORATE SOURCE: Turk. NII Glazn. Bolezn., Ashkhabad, USSR  
SOURCE: Zdravookhranenie Turkmenistana (1990), (10), 15-18  
CODEN: ZDTUAB; ISSN: 0513-8736  
DOCUMENT TYPE: Journal  
LANGUAGE: Russian

- AB The histochem. effects of 2-3-wk course treatment with histaglobin (histamine-binding globulin) on eye conjunctiva were studied in humans suffering from spring allergic conjunctivitis. The treatment decreased infiltration with basophils responsible for amine release and allergic symptoms.
- CC 1-7 (Pharmacology)  
Section cross-reference(s): 15
- ST allergic conjunctivitis histochem histaglobulin histamine binding;  
**histamine binding protein** allergic conjunctivitis histochem
- IT Proteins, specific or class  
RL: BIOL (Biological study)  
(histamine-binding, allergic spring conjunctivitis histochem. response to treatment with, in humans)
- IT Eye, disease or disorder  
(vernal conjunctivitis, histaglobin treatment effects on histochem. of, in humans)

L32 ANSWER 8 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1981:404871 HCAPLUS  
DOCUMENT NUMBER: 95:4871  
TITLE: **Histamine binding proteins**  
separated from human serums by the chromatographic method  
AUTHOR(S): Chachaj, Wladyslaw; Bartecka, Zuzanna; Malolepszy, Jozef  
CORPORATE SOURCE: Dep. Intern. Dis., Inst. Intern. Dis., Wroclaw, Pol.  
SOURCE: Archivum Immunologiae et Therapiae Experimentalis (1980), 28(6), 947-51  
CODEN: AITEAT; ISSN: 0004-069X  
DOCUMENT TYPE: Journal  
LANGUAGE: English

- AB Human serum proteins were bound on and eluted from Sepharose-polylysine-histamine column. Thus obtained protein fraction was referred to as HBP (**histamine-binding protein**). HBPs were examd. for the ability to bind histamine by biol. method applied to isolated guinea pig intestine. Chromatog. sepns. were run on DEAE cellulose column to obtain HBPs which are directly responsible for histamine-binding

**THIS PAGE BLANK (USPTO)**



activity. The results suggest that human serum contains 3 HBP fractions. One of them was identified as orosomucoid, and the 2 remaining ones seem to be glycoproteins belonging to .alpha.1 globulin group.

CC 15-1 (Immunochemistry)  
ST histamine binding protein chromatog  
IT Blood serum  
(histamine-binding proteins of)  
IT Allergy  
(histamine-binding proteins of blood  
serum in relation to)  
IT 51-45-6, biological studies  
RL: BIOL (Biological study)  
(proteins binding, of blood serum)

L32 ANSWER 9 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1979:536384 HCAPLUS  
DOCUMENT NUMBER: 91:136384  
TITLE: Application of affinity chromatography in binding of  
histamine to serum proteins  
AUTHOR(S): Bartecka, Z.; Chachaj, W.; Malepszy, J.  
CORPORATE SOURCE: Inst. Intern. Med., Med. Acad. Wroclaw, Wroclaw, Pol.  
SOURCE: Agents and Actions (1979), 9(1), 35-7  
CODEN: AGACBH; ISSN: 0065-4299  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB Serum histamine (I)-binding proteins (HBP) were isolated by affinity chromatog. on polylysine-Sepharose 4B-immobilized histamine (II) and characterized by electrophoresis. Dialyzed serum proteins were incubated with a suspension of carrier II for 2 h at 37.degree., and the carrier then was packed in a column, washed with 0.9% NaCl to remove unbound proteins, and eluted with 0.2M EDTA and 6M urea to remove HBP. Polyacrylamide gel electrophoresis in the presence of 0.2% Na dodecyl sulfate sepd. 8 HBP fractions with mol. wts. in the range 80,000-350,000. A bioassay with isolated guinea pig intestine was used to det. I and to study its inactivation at various HBP concns.

CC 9-2 (Biochemical Methods)  
ST serum histamine binding protein sepn;  
affinity chromatog histamine binding protein  
IT Blood analysis  
(histamine-binding proteins affinity.  
chromatog. in)  
IT Chromatography, column and liquid  
(affinity, of histamine-binding proteins)  
IT Proteins  
RL: PROC (Process)  
(histamine-binding, affinity chromatog. of, of blood serum)  
IT 9012-36-6D, reaction product with poly-L-lysine  
RL: ANST (Analytical study)  
(histamine immobilization on, for histamine-binding  
proteins affinity chromatog.)  
IT 51-45-6, biological studies  
RL: BIOL (Biological study)  
(proteins binding, affinity chromatog. of, of blood serum)

L32 ANSWER 10 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1978:35352 HCAPLUS  
DOCUMENT NUMBER: 88:35352  
TITLE: Histamine and serotonin content in patients with newly  
identified tuberculosis of the lungs complicated by  
nonspecific endobronchitis

**THIS PAGE BLANK (USPTO)**

AUTHOR(S): Shesterina, M. V.; Mikhailova, Yu. V.; Andrzheyuk, N. I.; Elistratova, N. A.  
 CORPORATE SOURCE: Mosk. Nauchno-Issled. Inst. Tuberk., Moscow, USSR  
 SOURCE: Problemy Tuberkuleza (1977), (9), 36-9  
 CODEN: PRTUAX; ISSN: 0032-9533  
 DOCUMENT TYPE: Journal  
 LANGUAGE: Russian

AB In 50 adult patients with tuberculosis and nonspecific endobronchitis, the blood levels of histamine and serotonin increased and those of **histamine-binding protein** decreased 2-, 1.5- and 5-fold, resp., in the patients without microflora in the bronchi and 1.4-, 1.5-, and 2-fold, resp., in the patients with addnl. microflora in the bronchi as compared with those of normal subjects. After treatment with antibiotics and antihistamines or corticosteroids, the levels usually normalized.

CC 14-1 (Mammalian Pathological Biochemistry)

ST histamine serotonin blood tuberculosis; bronchitis microorganism  
 tuberculosis; protein serum histamine binding endobronchitis

IT Tuberculosis

(blood compn. in, endobronchitis in relation to)

IT Proteins

RL: BIOL (Biological study)

(histamine-binding, of blood in tuberculosis and endobronchitis)

IT Bronchi

(disease, endobronchitis, blood compn. in tuberculosis in relation to)

IT 50-67-9, biological studies 51-45-6, biological studies

RL: BIOL (Biological study)

(of blood, in tuberculosis and endobronchitis)

L32 ANSWER 11 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1975:81859 HCAPLUS

DOCUMENT NUMBER: 82:81859

TITLE: Histamine-serum protein bond

AUTHOR(S): Kochman, S.; Jalouzet, R.; Maujean, A.; Cazabat, A.;  
 Dubois de Montreynaud, J. M.

CORPORATE SOURCE: Serv. Pneumophysiologie, Cent. Hosp. Univ., Reims, Fr.  
 SOURCE: Nouvelle Presse Medicale (1974), 3(30), 1883

CODEN: NPMDAD; ISSN: 0301-1518

DOCUMENT TYPE: Journal

LANGUAGE: French

AB Gel chromatog. of normal human serum on a histamine-fixed, CNBr-activated Sepharose 4B column adsorbed the histaminoplexig compd. of the serum. Elution with 6M guanidine-HCl yielded a protein (max. uv absorbance at 280 nm). Na dodecyl sulfate-polyacrylamide gel electrophoresis showed only 2 bands. Thus, the presence of a **histamine-binding protein** in blood serum is confirmed.

CC 6-3 (General Biochemistry)

ST blood serum protein histamine binding

IT Proteins

RL: BIOL (Biological study)

(blood-serum, histamine binding by)

IT 51-45-6, biological studies

RL: BIOL (Biological study)

(protein binding, of blood serum)

L32 ANSWER 12 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1974:502483 HCAPLUS

DOCUMENT NUMBER: 81:102483

TITLE: New aspects of histaminopexy

AUTHOR(S): Lonovics, J.; Gecse, A.; Karady, S.

**THIS PAGE BLANK (USPTO)**

CORPORATE SOURCE: Inst. Pathophysiol., Univ. Sch. Med., Szeged, Hung.  
 SOURCE: Archives Internationales de Physiologie et de  
 Biochimie (1974), 82(1), 41-7  
 CODEN: AIPBAY; ISSN: 0003-9799

DOCUMENT TYPE: Journal  
 LANGUAGE: English

AB The mechanism of formation and release of histaminopexic peptide was studied in the rat. Detns. were carried out with plasma and serum which were first subjected to gel filtration on Sephadex G-25. Agarose gel electrophoresis was used to sep. polypeptides, and free histamine (I) was assayed with an atropinized guinea pig ileum suspended in Tyrode's soln. Whereas I-binding activity of plasma was in the protein fraction, in serum it was in the peptide region; after papain digestion, results with plasma were similar to those with serum. Adrenalectomy caused loss of I-binding activity, unless rats were treated with prednisolone succinate. The I-binding peptide probably released from albumin by a papainlike action during blood coagulation. Adrenal cortex function is needed for maintenance of I-binding activity.

CC 13-13 (Mammalian Biochemistry)

ST histamine binding protein blood; adrenal  
 histamine binding protein

IT Proteins

RL: BIOL (Biological study)

(histamine binding, blood-plasma, regulation of)

IT Corticosteroids, biological studies

RL: BIOL (Biological study)

(histamine-binding peptide of blood serum in response to)

IT Peptides, biological studies

RL: BIOL (Biological study)

(histamine-binding, of blood serum, regulation of)

IT 2920-86-7

RL: BIOL (Biological study)

(histamine-binding peptide of blood serum in response to)

IT 51-45-6, biological studies

RL: BIOL (Biological study)

(peptide of blood serum binding of, regulation of)

L32 ANSWER 13 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1972:111092 HCAPLUS

DOCUMENT NUMBER: 76:111092

TITLE: Histamine-binding property of human serum

AUTHOR(S): Gecse, A.; Lonovics, J.; Szekeres, L.; Zsilinszky, E.;  
 West, G. B.

CORPORATE SOURCE: Sch. Med., Univ. Szeged, Szeged, Hung.

SOURCE: Journal of Pharmacy and Pharmacology (1972), 24(1),  
 70-1

CODEN: JPPMAB; ISSN: 0022-3573

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Expts. performed on serum from healthy adult humans indicated that the albumin fraction in some way binds histamine, whereas all other fractions of the plasma were lacking in such activity. When the fraction contg. 96% of albumin was digested with papain, and then passed through Sephadex G-25, the histamine-binding activity moved to the polypeptide region of the plasma. All the low-mol.-wt. polypeptide portion had been released from the albumin during coagulation, or by papain digestion. Plasma albumin or serum peptide from blood of 8 allergic patients (generalized dermatoses) lacked histamine-binding activity, indicating that lack of a histaminopexic substance may be involved in the pathogenesis of allergic diseases.

**THIS PAGE BLANK (USPTO)**

CC 13 (Mammalian Biochemistry)  
 ST blood histamine binding protein  
 IT Allergy  
     (histamine binding by albumins of blood serum in)  
 IT Albumins, blood serum  
     RL: BIOL (Biological study)  
     (histamine binding by, in allergy)  
 IT 51-45-6, biological studies  
     RL: BIOL (Biological study)  
     (albumins of blood serum binding of, in allergy)

L32 ANSWER 14 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1968:75101 HCAPLUS

DOCUMENT NUMBER: 68:75101

TITLE: Specificity effects of cation binding to the macroion  
     in the bovine serum albumin-histamine interaction

AUTHOR(S): Botre, Claudio; Marchetti, Marcello; Borghi, Saverio

CORPORATE SOURCE: Univ. Rome, Rome, Italy

SOURCE: Biochimica et Biophysica Acta (1968), 154(2), 360-6

CODEN: BBACAQ; ISSN: 0006-3002

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The specific interaction between bovine serum albumin and histamine and the influence of the ionic medium on this interaction were studied. This interaction is dependent on ionic strength. Evidence is given for a crit. saline concn. which can be regarded as a threshold defining 2 sharply different behaviors of bovine serum albumin in the presence of histamine. Potentiometric measurements carried out by means of coupled perm-selective membrane electrodes provided data which showed that for a given ionic strength, bovine serum albumin requires an amt. of histamine which varies with the nature of the cations. Alkali binding to the bovine serum albumin macroion in the presence of histamine does not follow the sequence expected for polycarboxylic acids, since K<sup>+</sup> proved to be the most tightly bound. 20 references.

CC 2 (General Biochemistry)

ST BINDING; HISTAMINE BINDING PROTEINS;  
 ELECTROLYTES ALBUMINS HISTAMINE; CATIONS ALBUMINS HISTAMINE BINDING;  
 ALBUMINS HISTAMINE BINDING; PROTEINS HISTAMINE BINDING

IT Cations, biological studies  
     (in histamine reaction with albumins)

IT Albumins, blood serum  
     RL: RCT (Reactant); RACT (Reactant or reagent)  
     (reaction of, with histamine, cation binding effect on)

IT 51-45-6, reactions  
     RL: RCT (Reactant); RACT (Reactant or reagent)  
     (with albumin, cation binding effect on)

**THIS PAGE BLANK (USPTO)**





XX Claim 4; Page 5-6; 19pp; English.  
 PS The present invention relates to the use of a histacalin protein  
 CC for treating or preventing conjunctivitis. The present sequence  
 CC is the histacalin protein FS-HBP1. The invention is particularly useful  
 CC in the treatment of allergic or seasonal conjunctivitis.  
 XX  
 SQ Sequence 172 AA;  
 Query Match 100.0%; Score 952; DB 22; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-94;  
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DKPWADAEANGEHQDAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNLNDE 60  
 DB 1 DKPWADAEANGEHQDAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNLNDE 60  
 QY 61 KNEWAFMFNNADTVYQHTEFEKATPDKMGYNKENAITYQTEGQVLTDLVLAESDNCY 120  
 DB 61 KNEWAFMFNNADTVYQHTEFEKATPDKMGYNKENAITYQTEGQVLTDLVLAESDNCY 120  
 QY 121 VYVALGPDGSGAGYELMATDVTDPASCLEKFNKYAGLPVRYVTSDCLEPE 172  
 DB 121 VYVALGPDGSGAGYELMATDVTDPASCLEKFNKYAGLPVRYVTSDCLEPE 172  
 RESULT 2  
 AAB73260  
 ID AAB73260 standard; protein; 172 AA.  
 AC AAB73260;  
 XX  
 DT 15-MAY-2001 (first entry)  
 XX  
 DE Histacalin protein FS-HBP1.  
 XX  
 KW FS-HBP1; histacalin protein; antinflammatory; antiallergic;  
 KM ophthalmological; allergic rhinitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200116164-A2.  
 XX  
 PD 08-MAR-2001.  
 XX  
 PF 24-AUG-2000; 2000WO-GB03287.  
 XX  
 PR 01-SEP-1999; 99GB-0020673.  
 XX  
 PA (EVOL-) EVOLUTEC LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 XX  
 DR WPI; 2001-218521/22.  
 XX  
 PT Use of histacalin proteins for treating or preventing allergic  
 PT rhinitis, or for manufacturing a medicament for treating or preventing  
 PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis  
 XX  
 PS Disclosure; Pages 4-6; 19pp; English.  
 XX  
 CC The present invention relates to a method for treating or preventing  
 CC allergic rhinitis. The method involves employing a blood-feeding  
 CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
 CC sequence is one such histacalin protein. The histacalin protein, is  
 CC useful for treating or preventing allergic rhinitis, both seasonal and  
 CC perennial allergic conjunctivitis.  
 XX  
 SQ Sequence 172 AA;  
 Query Match 100.0%; Score 952; DB 22; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-94;

Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DKPWADAEANGEHQDAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNLNDE 60  
 DB 1 DKPWADAEANGEHQDAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNLNDE 60  
 QY 61 KNEWAFMFNNADTVYQHTEFEKATPDKMGYNKENAITYQTEGQVLTDLVLAESDNCY 120  
 DB 61 KNEWAFMFNNADTVYQHTEFEKATPDKMGYNKENAITYQTEGQVLTDLVLAESDNCY 120  
 QY 121 VYVALGPDGSGAGYELMATDVTDPASCLEKFNKYAGLPVRYVTSDCLEPE 172  
 DB 121 VYVALGPDGSGAGYELMATDVTDPASCLEKFNKYAGLPVRYVTSDCLEPE 172  
 RESULT 3  
 AAW37446  
 ID AAW37446 standard; Protein; 190 AA.  
 AC AAW37446;  
 XX  
 DT 08-JUN-1998 (first entry)  
 XX  
 DE Tick vasoactive amine binding protein 1 FS-HBP1.  
 XX  
 KW Female-specific vasoactive amine binding protein 1; FS-HCP1;  
 KW histamine; serotonin; assay; antihistamine; anti-inflammatory;  
 KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
 KW transgenic animal; tick.  
 XX  
 OS Rhipicephalus appendiculatus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..18  
 FT /label= sig\_peptide  
 PN WO9744451-A2.  
 XX  
 PD 27-NOV-1997.  
 XX  
 PF 19-MAY-1997; 97WO-GB01372.  
 XX  
 PR 18-APR-1997; 97GB-0007844.  
 PR 18-MAY-1996; 96GB-0010484.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 XX  
 DR WPI; 1998-018506/02.  
 DR N-PSDB; AAV00227.  
 XX  
 PT New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 XX  
 PS Example 2; Fig 1; 44pp; English.  
 XX  
 CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
 CC female-specific histamine binding protein 1 (FS-HBP1). Its amino  
 CC acid sequence was deduced from a cDNA clone (see AAV00227) obtained  
 CC from a salivary gland cDNA library. FS-HBP2 and male-specific  
 CC HBP1 (see AAW37447-48) and a related protein, D-RET6 (see AAW37449)  
 CC from Dermacentor reticulatus, were also identified. These novel  
 CC vasoactive amine binding proteins (VABPs) can be expressed in  
 CC host cells using e.g. a baculovirus expression system. They can  
 CC be used: (i) to assay histamine (or other VA such as serotonin) in  
 CC body fluids or cell culture supernatants, e.g. to monitor the  
 CC effect of allergens; (ii) for binding VA, e.g. to remove histamine  
 CC from blood, food, cell cultures etc.; (iii) as an antihistamine or  
 CC anti-inflammatory agents, e.g. for treating insect, snake or  
 CC scorpion bites or dermatitis, or as a carrier for slow release of

CC histamine-related compounds; (iv) in vaccines to protect against  
CC metazoan parasites, especially in animals; (v) as reagents for the  
CC studying inflammation, involvement of VA in ulcer formation or the  
CC immune response etc. VABPs provide a more sensitive assay for  
CC histamine than low-affinity antibodies currently used. They may  
CC also be more effective and safer than conventional antihistamines.

SO Sequence 190 AA;

Query Match 100.0%; Score 952; DB 19; Length 190;  
Best Local Similarity 100.0%; Pred. No. 4,5e-94;  
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKPVMDEANGEHQDAMKHLQKLYEENYDLKATYKNDPVWGNDPFCVGTAAQNLMDE 60  
DB 19 DKPVMDEANGEHQDAMKHLQKLYEENYDLKATYKNDPVWGNDPFCVGTAAQNLMDE 78  
QY 61 KNEVAMFPMNNADTVYQHTEFEKATPKMYGYNKENAITYOTEDGQVLTDLVLAESDNCY 120  
DB 79 KNEVAMFPMNNADTVYQHTEFEKATPKMYGYNKENAITYOTEDGQVLTDLVLAESDNCY 138  
QY 121 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 172  
DB 139 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 190

## RESULT 4

AAV18078  
ID AAV18078 standard; protein: 190 AA.

AAV18078;

DT 06-AUG-1999 (first entry)

DE Histamine binding protein FS-HBP1.

XX Histamine binding protein; serotonin binding compound; inflammation;  
KM gastric acid secretion; allergy; type I hypersensitivity reaction;  
KM asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KM drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KM respiratory disease; coronary heart disease; cellular growth regulator;  
KM tissue repair; blood-sucking ectoparasite; therapy.

OS Rhinolephalus appendiculatus.

XX W09927104-A1.

PD 03-JUN-1999.

XX 26-NOV-1998; 98MO-G803530.

XX 26-JUN-1998; 98GB-0013917.

XX 26-NOV-1997; 97GB-0025046.

PA (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC;

DR WPI: 1999-357841/30.

XX N-PSDB; AAX76964.

XX Histamine and serotonin binding compounds useful for the treatment  
PT of allergies

XX Claim 1; Fig 1; 84pp; English.

CC This sequence is an example of a histamine or serotonin binding  
CC compound (A), of the invention. The compounds are useful for regulating  
CC the action of histamine and serotonin (in e.g. inflammation and gastric  
CC acid secretion), the detection, quantification and removal of histamine  
CC or serotonin (in animals, plants, cell cultures, food materials, or  
CC humans) and in the treatment of various diseases and allergies  
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic

CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
CC allergies, abnormal blood pressure, migraine, psychological disorders,  
CC respiratory disease, and coronary heart disease). Histamine may also be  
CC used to regulate cellular growth and tissue repair. The molecules may  
CC also be used as components of vaccines directed against blood-sucking  
CC ectoparasites.

SO Sequence 190 AA;

Query Match 100.0%; Score 952; DB 20; Length 190;  
Best Local Similarity 100.0%; Pred. No. 4,5e-94;  
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKPVMDEANGEHQDAMKHLQKLYEENYDLKATYKNDPVWGNDPFCVGTAAQNLMDE 60  
DB 19 DKPVMDEANGEHQDAMKHLQKLYEENYDLKATYKNDPVWGNDPFCVGTAAQNLMDE 78  
QY 61 KNEVAMFPMNNADTVYQHTEFEKATPKMYGYNKENAITYOTEDGQVLTDLVLAESDNCY 120  
DB 79 KNEVAMFPMNNADTVYQHTEFEKATPKMYGYNKENAITYOTEDGQVLTDLVLAESDNCY 138  
QY 121 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 172  
DB 139 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 190

## RESULT 5

AAB74289  
ID AAB74289 standard; protein: 171 AA.

AAB74289;

DT 20-JUN-2001 (first entry)

DE Histacalin protein FS-HBP2.

XX Histacalin; FS-HBP1; conjunctivitis.

XX Unidentified.

XX W0200115719-A2.

XX 08-MAR-2001.

XX 24-AUG-2000; 2000MO-G803282.

XX 01-SEP-1999; 99GB-0020674.

PA (EVOL-) EVOLUTEC LTD.

XX Nuttall PA, Paesen GC;

DR WPI: 2001-257675/26.

XX Use of histacalin proteins for treating or preventing non-infective  
PT conjunctivitis, or for manufacturing a medicament for treating or  
PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
PT conjunctivitis

XX Claim 4; Page 5-6; 19pp; English.

CC The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein FS-HBP1. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.

XX Sequence 171 AA;

Query Match 66.2%; Score 630; DB 22; Length 171;  
Best Local Similarity 66.5%; Pred. No. 1,7e-59;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 DKPVMDEANGEHQDAMKHLQKLYEENYDLKATYKNDPVWGNDPFCVGTAAQNLMDE 60





```

XX AC AAB73262;
XX DE 15-MAY-2001 (first entry)
XX DE Histacalin protein MS-HBPI.
XX MS-HBPI; histacalin protein; antiinflammatory; antiallergic;
XX opthalmological; allergic rhinitis.
XX OS Unidentified.
XX FN W0200116164-A2.
XX PD 08-MAR-2001.
XX PF 24-AUG-2000; 2000WO-GB03287.
XX PR 01-SEP-1999; 99GB-0020673.
XX PA (EVOL-) EVOLUTEC LTD.
XX PI Nuttall PA, Paesen GC;
XX DR WPI; 2001-218521/22.
XX PT Use of histacalin proteins for treating or preventing allergic
XX rhinitis, or for manufacturing a medicament for treating or preventing
XX allergic rhinitis, e.g. seasonal or perennial allergic rhinitis.
XX PS Disclosure; Pages 4-6; 19pp; English.
XX CC The present invention relates to a method for treating or preventing
XX allergic rhinitis. The method involves employing a blood-feeding
XX ectoparasite-derived (e.g. tick-derived) histacalin protein. The present
XX sequence is one such histacalin protein. The histacalin protein, is
XX useful for treating or preventing allergic rhinitis, both seasonal and
XX perennial allergic conjunctivitis.
XX SQ Sequence 182 AA;

Query Match 28.1%; Score 267.5; DB 22; Length 182;
Best Local Similarity 39.3%; Pred. No. 1.9e-20;
Matches 70; Conservative 20; Mismatches 77; Indels 11; Gaps 7;

OY 3 PWADEAANGHODAMKHLQKLYEENYDLKATYKNDPVNGNDFTCVGTAQNLNDEKN 62
DB 2 PTWANEAKLSYQDAMKSLQDQDKRYLAQAQTGDSVGEFTCVSYAEKIG--KKR 59
OY 63 VEAMFEMNNADPVYQHTFEKATPDKMYGYNKENATYQTED--GOVLTVLAFSD-DNC 119
DB 60 LNAFTILYKKNHLDLAKESHETITVWKAAYDTTENGKIKYETQGRTRQTFEDVFESDYKNC 119
OY 120 YVITVYALGPDGSGAG-YELMAT-DYTDVPASCLEKFNENYAA---GLPVDVYT--SDCLP 171
DB 120 DVIFVPERKERSDEGDELMWSEDKIDIPCCKFTMAFYAQOQEKTVRNYYTDSCKP 177

RESULT 11
ID AAM37448 standard; Protein; 200 AA.
XX AAM37448;
XX 08-JUN-1998 (first entry)
XX
XX Tick vasoactive amine binding protein 1 MS-HBPI.
XX Male-specific vasoactive amine binding protein 1; MS-HBPI;
XX histamine; serotonin; assay; antihistamine; anti-inflammatory;
XX insect bite; snake bite; scorpion bite; dermatitis; vaccine;
XX transgenic animal; tick.

```

```

OS Rhipicephalus appendiculatus.
XX FH Key Location/Qualifiers
XX FT Peptide 1..18
XX FT /label= Sig-peptide
XX FT Modified-site 79..81
XX FT /note= "Asn is N-glycosylated"
XX FN W09744451-A2.
XX PD 27-NOV-1997.
XX PF 19-MAY-1997; 97WO-GB01372.
XX PR 18-APR-1997; 97GB-0007844.
XX PR 18-MAY-1996; 96GB-0010484.
XX PA (OXFO-) OXFORD VACS LTD.
XX PI Nuttall PA, Paesen GC;
XX DR WPI; 1998-018506/02.
XX DR N-PSDB; AAV00229.
XX PT New vasoactive amine binding proteins and related nucleic acid,
XX vectors - transformed cells and transgenic animals, used for
XX assaying or removing histamine and as antihistamine or
XX anti-inflammatory agents
XX PS Example 2; Fig 3; 44pp; English.
XX CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel
XX male-specific histamine binding protein 1 (MS-HBPI). Its amino
XX acid sequence was deduced from a cDNA clone (see AAV00227) obtained
XX from a salivary gland cDNA library. Female-specific HSP1 and
XX HSP2 (see AAM37446-47) and a related protein, D.RB76 (see AAM37449)
XX from Dermacentor reticulatus, were also identified. These novel
XX vasoactive amine binding proteins (VABPs) can be expressed in
XX host cells using e.g. a baculovirus expression system. They can
XX be used: (i) to assay histamine (or other VA such as serotonin) in
XX body fluids or cell culture supernatants, e.g. to monitor the
XX effect of allergens; (ii) for binding VA, e.g. to remove histamine
XX from blood, food, cell cultures etc.; (iii) as an antihistamine or
XX anti-inflammatory agents, e.g. for treating insect, snake or
XX scorpion bites or dermatitis; or as a carrier for slow release of
XX histamine-related compounds; (iv) in vaccines to protect against
XX metazoan parasites, especially in animals; (v) as reagents for
XX studying inflammation, involvement of VA in ulcer formation for the
XX immune response etc. VABPs provide a more sensitive assay for
XX histamine than low-affinity antibodies currently used. They may
XX also be more effective and safer than conventional antihistamines.
XX SQ Sequence 200 AA;

Query Match 28.1%; Score 267.5; DB 19; Length 200;
Best Local Similarity 39.3%; Pred. No. 2.1e-20;
Matches 70; Conservative 20; Mismatches 77; Indels 11; Gaps 7;

OY 3 PWADEAANGHODAMKHLQKLYEENYDLKATYKNDPVNGNDFTCVGTAQNLNDEKN 62
DB 20 PTWANEAKLSYQDAMKSLQDQDKRYLAQAQTGDSVGEFTCVSYAEKIG--KKR 77
OY 63 VEAMFEMNNADPVYQHTFEKATPDKMYGYNKENATYQTED--GOVLTVLAFSD-DNC 119
DB 78 LNAFTILYKKNHLDLAKESHETITVWKAAYDTTENGKIKYETQGRTRQTFEDVFESDYKNC 137
OY 120 YVITVYALGPDGSGAG-YELMAT-DYTDVPASCLEKFNENYAA---GLPVDVYT--SDCLP 171
DB 138 DVIFVPERKERSDEGDELMWSEDKIDIPCCKFTMAFYAQOQEKTVRNYYTDSCKP 195

RESULT 12
AAV18080

```

ID	AA118080 standard; Protein: 200 AA.
XX	AA118080;
AC	06-AUG-1999 (first entry)
DT	Histamine binding protein MS-HBPL
XX	
DE	Histamine binding protein; serotonin binding compound; inflammation;
KW	gastric acid secretion; allergy; type I hypersensitivity reaction;
KW	asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW	drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW	respiratory disease; coronary heart disease; cellular growth regulator;
KW	tissue repair; blood-sucking ectoparasite; therapy.
OS	Rhipicephalus appendiculatus.
XX	
PN	MO9927104-A1.
XX	
PD	03-JUN-1999.
XX	
PE	26-NOV-1998; 98WO-GB03530.
XX	
PR	26-JUN-1998; 98GB-0013917.
PR	26-NOV-1997; 97GB-0025046.
XX	
PA	(OXFO-) OXFORD VACS LTD.
XX	
PI	Nuttall PA, Paesen GC;
DR	WPI: 1999-357841/30.
XX	
XX	N-PSDB: AAX76966.
XX	
PT	Histamine and serotonin binding compounds useful for the treatment
XX	of allergies
PS	Claim 1; Fig 3; 84pp: English.
XX	
CC	This sequence is an example of a histamine or serotonin binding
CC	compound (A), of the invention. The compounds are useful for regulating
CC	the action of histamine and serotonin (in e.g. inflammation and gastric
CC	acid secretion), the detection, quantification and removal of histamine
CC	or serotonin (in animals, plants, cell cultures, food materials, or
CC	humans) and in the treatment of various diseases and allergies
CC	(e.g. type I hypersensitivity reactions, urticaria/asthma, allergic
CC	rhinitis (hay fever), atopic dermatitis, insect bites and food and drug
CC	allergies, abnormal blood pressure, migraine, psychological disorders,
CC	respiratory disease, and coronary heart disease). Histamine may also be
CC	used to regulate cellular growth and tissue repair. The molecules may
CC	also be used as components of vaccines directed against blood-sucking
CC	ectoparasites.
SO	
Sequence	200 AA:
Query Match	28.1%; Score 267.5; DB 20; Length 200;
Best Local Similarity	39.3%; Pred. No. 2.1e-20;
Matches	70; Conservative 20; Mismatches 77; Indels 11; Gaps 7.
OY	3 PWADEANGEHODAKHLOKLEVENYDLIKATYKNDPVWGNDFCVCTAONLNEDKN 62
DB	20 PFWANEARLGSTQDAWMSKLODQONRRYLAQNTOTTDVWGSEFVSVLTAEKII--KKK 77
OY	63 VEAMFPMNNADTVYQHTFEKATPDKMYGKNENATVQTED--GQVLTDLVAFSD-DNC 119
DB	78 LNAITLYKNKHILTDKESHETITVWKAADYTTENKIKKETQRTQRTQEDVFEFDYKNC 137
OY	120 VYIYALGPDGSGAG-YELMAT-DYTDVPASCLEKENEYAA---GLPVDVYT-SDCLP 171
DB	138 DVIEFVPRKRGSGDEGDEYELWVSEDKIDKIPDCCKFTMAFYAFOOQKERTVNRVYTDSSCKP 195

ID AAW37449 standard; Protein: 209 AA.  
XX  
XX AAW37449;  
AC  
XX  
DT 08-JUN-1998 (first entry)  
XX  
DE Tick vasoactive amine binding protein D.RET6.  
XX  
XX Vasoactive amine binding protein; D.RET6; histamine; serotonin;  
KW assay; antihistamine; anti-inflammatory; insect bite; snake bite;  
KW scorpion bite; dermatitis; vaccine; transgenic animal; tick.  
XX  
OS Dermacenter reticularis.

Key Location/Qualifiers  
FH Peptide 1..28  
FT /label= sig\_peptide

MW0974451-AZ.  
XX  
XX PD 27-NOV-1997.  
XX PF 19-MAY-1997; 97WO-GB01372.  
XX PR 18-APR-1997; 97GB-0007844.  
PR 18-MAY-1996; 96GB-0010484.  
PA (OXFO-) OXFORD VACS LTD.  
XX  
PI Nuttall PA, Paesen GC;  
DR WPI: 1998-018506/02.  
XX N-PDSB: AAV00230.

New vasoactive amine binding proteins and related nucleic acid,  
PT vectors - transformed cells and transgenic animals, used for  
PT assaying or removing histamine and as antihistamine or  
PT anti-inflammatory agents

Example 2; Fig 4; 44pp; English.

This protein comprises tick Dermacenter reticularis (Dr) novel  
CC vasoactive amine binding protein (VABP) D.RET6, its amino acid  
CC sequence was deduced from a cDNA clone (see AAV00230) obtained  
CC from a Dr salivary gland cDNA library. 3 Novel VASPs, designated  
CC FS-HBP1, FS-HBP2 and MS-HBP1 (see AAW37446-48), of the tick  
CC Rhipicephalus appendiculatus have also been identified. The  
CC VABPs can be expressed in host cells using e.g. a baculovirus  
CC expression system. They can be used: (i) to assay histamine (or  
CC other VA such as serotonin) in body fluids or cell culture  
CC supernatants, e.g. to monitor the effect of allergens; (ii) for  
CC binding VA, e.g. to remove histamine from blood, food, cell  
CC cultures etc.; (iii) as an antihistamine or anti-inflammatory  
CC agents, e.g. for treating insect, snake or scorpion bites or  
CC dermatitis, or as a carrier for slow release of histamine-related  
CC compounds; (iv) in vaccines to protect against metazoan parasites,  
CC especially in animals; (v) as reagents for studying inflammation,  
CC involvement of VA in ulcer formation or the immune response etc.  
CC VABPs provide a more sensitive assay for histamine than  
CC low-affinity antibodies currently used. They may also be more  
CC effective and safer than conventional antihistamines.

Sequence 209 AA;

Query Match 25.1%; Score 239; DB 19; Length 209;  
Best Local Similarity 34.5%; Pred. No. 2.6e-17;  
Matches 61; Conservative 24; Mismatches 82; Indels 10; Gaps 7

1 DKPVADDAANGEHODAMKHLQLEEVYDILIKATYYNDP-VGNDPFCTGTAAQNLED 59  
|::|| |:::||| |::| |::|| ||::||  
Db ENPLMAHELLGGTKVDAMKSIDCGSVTYVLAKTTIENDPSMGSGFCLGOVOELKEEE 88  
60 EKNEVAEMFNMMNADV-YQHTEFKARPDKNKY-NKENAITYTQTEDGVLTDLVASFD- 116

Db 89 DYTWTSEFTFRNASSPIKYYNVTETKAVFOYGYKKNIRNAIEYOVGGGLNITDPLIFTDG 148  
QY 117 DNCVYIALGPDGSGAGYEELMA--TDYTDVPASGLEFNEFYAA-GLFVRVYISDCL 170  
Db 149 ELCDVFYVPNAD--QGCELMVKKSHYKHVPDICTFEVNFVCAKDRRTYDIFNEECV 202

## RESULT 14

AAV18081  
ID AAV18081 standard; Protein; 209 AA.

AAV18081;

06-AUG-1999 (first entry)

Histamine binding protein D.RET6.

KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.

Rhipicephalus appendiculatus.

MO9927104-A1.

03-JUN-1999.

26-NOV-1998; 98MO-GB03530.

26-JUN-1998; 98GB-0013917.

26-NOV-1997; 97GB-0025046.

(OXFO-) OXFORD VACS LTD.

Nuttall PA, Paesen GC;

WPI; 1999-357841/30.

N-PSDB; AAX76967.

Histamine and serotonin binding compounds useful for the treatment of allergies

Claim 1; Fig 4; 84pp; English.

This sequence is an example of a histamine or serotonin binding compound (A), of the invention. The compounds are useful for regulating the action of histamine and serotonin (in e.g. inflammation and gastric acid secretion), the detection, quantification and removal of histamine or serotonin (in animals, plants, cell cultures, food materials, or humans) and in the treatment of various diseases and allergies (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis, insect bites and food and drug allergies, abnormal blood pressure, migraine, psychological disorders, respiratory disease, and coronary heart disease). Histamine may also be used to regulate cellular growth and tissue repair. The molecules may also be used as components of vaccines directed against blood-sucking ectoparasites.

Sequence 209 AA;

Query Match 25.1%; Score 239; DB 20; Length 209;

Best Local Similarity 34.5%; Pred. NO.2.6e-17; Indels 10; Gaps 7;

Matches 61; Conservative 24; Mismatches 82;

QY 1 DKNVMADEANGEHODAMKHLQKLVENYDLIKATYKNDP-VGNGDFTGVTAAQNIINED 59

Db 29 ENPLAMHEELGTYQAMKSIOGVSVTYLAKTTEENDTSGSFKCLQVGEIERKEE 88

QY 60 EKNVEMFMNNAATV-YQHTFERKATPDKMIGY-NKENAIYQTEDEGVLDVLAFSO- 116

Db 89 DYTWTSEFTFRNASSPIKYYNVTETKAVFOYGYKKNIRNAIEYOVGGGLNITDPLIFTDG 148  
QY 117 DNCVYIALGPDGSGAGYEELMA--TDYTDVPASGLEFNEFYAA-GLFVRVYISDCL 170  
Db 149 ELCDVFYVPNAD--QGCELMVKKSHYKHVPDICTFEVNFVCAKDRRTYDIFNEECV 202

## RESULT 15

AAV18085  
ID AAV18085 standard; Protein; 203 AA.

AAV18085;

06-AUG-1999 (first entry)

Histamine binding protein Ih/Bm-HBPI.

KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.

Boophilus microplus.

Ixodes hexagonus.

Synthetic.

MO9927104-A1.

03-JUN-1999.

26-NOV-1998; 98MO-GB03530.

26-JUN-1998; 98GB-0013917.

26-NOV-1997; 97GB-0025046.

(OXFO-) OXFORD VACS LTD.

Nuttall PA, Paesen GC;

WPI; 1999-357841/30.

N-PSDB; AAX76970.

Histamine and serotonin binding compounds useful for the treatment of allergies

Claim 13; Fig 7; 84pp; English.

This sequence is an example of a histamine or serotonin binding compound (A), of the invention. cDNA encoding this sequence was isolated from a mixed Boophilus microplus/Ixodes hexagonus cDNA expression library. The compounds are useful for regulating the action of histamine and serotonin (in e.g. inflammation and gastric acid secretion), the detection, quantification and removal of histamine or serotonin (in animals, plants, cell cultures, food materials, or humans) and in the treatment of various diseases and allergies (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis, insect bites and food and drug allergies, abnormal blood pressure, migraine, psychological disorders, respiratory disease, and coronary heart disease). Histamine may also be used to regulate cellular growth and tissue repair. The molecules may also be used as components of vaccines directed against blood-sucking ectoparasites.

Sequence 203 AA;

Query Match 22.0%; Score 209.5; DB 20; Length 203;

Best Local Similarity 31.2%; Pred. No.3.8e-14; Indels 7; Gaps 5;

Matches 50; Conservative 26; Mismatches 75;

QY 14 HODAMKHLQKLVENYDLIKATYKNDPVGNGDFTGVTAAQNIINEDKENVEMFMNNA 73



Db 43 YODAKSINONVSTTYFFLSTYNNDSVWGKNFTCLSVTVTSKHSTFTVEYNTTKNOS 102  
QY 74 DTYYQHTEKATPDKNMGYKNENAITYQTEDEQVLTVDYLAESD-DNCTYIYALGPDGSGA 132  
Db 103 OQWVSMT-ENVTAVOEGYDVKNILITOMTTENNKKFNDTVVFTDGOCTCDLXIPYKEN-- 158  
QY 133 GYELMA-TDY-TDVPASCLEKFNENYAAGLPVRDYVTSACL 170  
159 GYELMVRSDYLONTPTCCQFIEDLVALGRRTYINISTPDCV 198

Search completed: August 1, 2003, 12:22:46  
Job time : 54.4019 secs

**THIS PAGE BLANK (USPTO)**



```

Db      DQ      DB      RESULT 2
79  NVEWMFMEHNNADIVYQHTEKAPDKMYGYGNKENALITYQIEDGVLTVDLAFSDNKC 138
121 VIALGPDGSGAGCYELMATDYTDVPASCLEKEKNEXAAGLPVADVTSDCPE 172
|||||
139 VIALGPDGSGAGCYELMATDYTDVPASCLEKEKNEXAAGLPVADVTSDCPE 190

Db      DQ      DB      RESULT 2
HBP2_RHIAP
HBP2_RHIAP STANDARD; PRT; 190 AA.
AC      077421;
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DE      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Female-specific histamine-binding protein 2 precursor (FS-HBP2).
OS      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC      Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.
OX      NCBI_TaxID=34631;
RN      [1]
RP      SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).
RC      TISSUE=Salivary gland;
RA      MEDLINE=99288454; PubMed=10360182;
RX      Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT      "Tick histamine-binding proteins: Isolation, cloning, and three-dimensional structure."
RL      Mol. Cell 3:661-671(1999).
CC      -I- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO
CC      OUTCOMPLETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO
CC      SUPPRESS INFLAMMATION DURING BLOOD FEEDING.
CC      -I- SUBCELLULAR LOCATION: Secreted.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation .
CC      CC the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U96081; AAC63107.1; -.
DR      PDB; 1OFT; 19-APR-00.
DR      PDB; 1OEV; 19-APR-00.
DR      InterPro; IPRO02970; His_binding.
DR      Pfam; PF02098; His_binding_1.
DR      ProDom; PD152455; His_binding_1.
FT      Signal, 3D-structure.
FT      SIGNAL          1           19
FT      CHAIN           20        190
FT                                     FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN
FT                                     2.
FT      DISULFID        67         188
FT      DISULFID       138        167
FT      TURN           23         24
FT      HELIX           27         30
FT      HELIX           31         33
FT      HELIX           36         41
FT      TURN           42         45
FT      STRAND          48         53
FT      STRAND          57         58
FT      TURN           59         61
FT      STRAND          62         63
FT      STRAND          64         64
FT      TURN           66         76
FT      STRAND          77         80
FT      TURN           81         89
FT      TURN           91         92
FT      STRAND          97         106
FT      TURN          109         110
FT      STRAND          116         121
FT      TURN          122         123
FT      STRAND          126         135
FT      TURN          136         137
FT      STRAND          138         143

```

```

FT STRAND 152 157
FT TURN 160 161
DT HELIX 165 174
FT TURN 175 177
FT STRAND 181 182
FT TURN 186 188
SQ SEQUENCE 190 AA; 21464 MW; 6923A3E902552B6F CRC64;

Query Match 66.2%; Score 630; DB 1; Length 190;
Best Local Similarity 66.5%; Pred. No. 1.3e-48;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

1 DKPWADNANGEHODAMKHLQKLEENYVDLIRKATYKNDPYWGNDFCTGVTAQNLNDE 60
20 NOPDADADANGAHODAMKSLKADENYVYWKATYKNDPYWGNDFCTGVWADVNDDE 79
61 KNVEAMFEMFNNAADTVYCHTEKATPDPMYGYKNENATYOTDEQVLTDLVAFSDNDCY 120
80 KSIQAEFLFNNADNTNMQFATEKAYVAMGYGNRENAPRYETEDQVFTVLYASDQCD 139
121 VYIAGPDGSGAGYELMATDYTDVDPASCLAKFNFNEYAACLPIYRDYTSCL 170
140 VYVPGTGTGNEGELYELMTDYTDNPANCLANFNEYAVAGRETRDFTVSACL 189

RESULT 3
HBPW_RHIAP STANDARD; PRT; 200 AA.
ID HBPW_RHIAP 077422;
DT 30-MAY-2000 (Rel. 39, created)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)
DE Male-specific histamine-binding salivary protein precursor (MS-HBP).
DS Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OX NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Salivary gland;
RX MEDLINE=99288454; PubMed=10360182;
RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT "tick histamine-binding proteins: isolation, cloning, and three-
dimensional structure.";
RL Mol. Cell 3:661-671(1999).
CC -I- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO
OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO
SUPPRESS INFLAMMATION DURING BLOOD FEEDING.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U96082; AAC63108.1; -.
DR HSSP: O77421; 10FT.
DR InterPro: IPR002870; His_binding.
DR Pfam: PF02098; His_binding_1.
DR ProDom: PD152455; His_binding; 1.
KW signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 200 MALE-SPECIFIC HISTAMINE-BINDING SALIVARY
FT PROTEIN.
FT DISULFID 65 193 BY SIMILARITY.
FT DISULFID 137 169 BY SIMILARITY.
SQ SEQUENCE 200 AA; 22851 MW; C46A1C08CBCA008 CRC64;

Query Match 28.1%; Score 267.5; DB 1; Length 200;
Best Local Similarity 39.3%; Pred. No. 1.1e-16;

```

APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-484  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2052 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-790-912-2

Query Match 8.5%; Score 81; DB 2; Length 2052;  
Best Local Similarity 24.1%; Pred. No. 15;  
Matches 45; Conservative 25; Mismatches 67; Indels 50; Gaps 8;

QY 12 GEHDDAKKHLOKLEEN-----YDLKATYKNDPPWGNDDTCVGTAAQNLNDEKNEA 65  
DB 1146 GTMLDASQIVSKAKELPLPLPTVEPLSTSGKD---SDFSKIAHYOANALVYKNIK 1201

QY 66 WFMFMNADYVYQHTFEKATPDKMYGYNKENATYOTE-----DGOVLTDLAFSD 117  
DB 1202 LLPRYN-----KSTIVKGNLVKENSILYOKELLSAVMMKDDQVTTDIVS-NKO 1249

QY 118 NCYVIVALGPDGSGAGYEL-WATDYTDVPASCLEKFNEYAG-----LPVRDYY 165  
DB 1250 TANLLHYNDHSEKFDLKYQTFANLP-----EYNIGNTGLTYTPNQFLYDRDSI 1301

QY 166 TSDCLPE 172  
DB 1302 VKEVLPE 1308

RESULT 5  
US-08-271-354-9  
Sequence 9, Application US/08271354  
Patent No. 5695983  
GENERAL INFORMATION:  
APPLICANT: Miller, Samuel I.  
APPLICANT: Melanos, John J.  
TITLE OF INVENTION: SALMONELLA VACCINES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/271,354  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/077,526  
FILING DATE: July 9, 1993  
APPLICATION NUMBER: 07/629,602

FILING DATE: December 18, 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/220001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-271-354-9

Query Match 8.4%; Score 80; DB 1; Length 246;  
Best Local Similarity 23.7%; Pred. No. 0.84;  
Matches 27; Conservative 21; Mismatches 46; Indels 20; Gaps 5;

QY 38 NDPVWGN---DFTCVGTAAQNLNDEKNEAWFMFMNADTV-----YQHTFEKAT 85  
DB 110 NNGYWGSPVNLTFGTITEE-OKOELKALPKF---DSINFSIIPERIOETIKRAN 164

QY 86 PDKMYGYNKENATYOTE---DGOVLTDLAFSDNCCVIVALGPDGSGAGYEL 136  
DB 165 ASGIISYEDSDIVRAEIAHNGSEYDITITANNTARAVTLLNKDGIAGYEI 218

RESULT 6  
US-08-565-861-9  
Sequence 9, Application US/08565861  
Patent No. 5843426  
GENERAL INFORMATION:  
APPLICANT: Miller, Samuel I.  
APPLICANT: Melanos, John J.  
APPLICANT: Hohmann, Elizabeth  
TITLE OF INVENTION: SALMONELLA VACCINES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/565,861  
FILING DATE: 01-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/271,354  
FILING DATE: 06-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/090,526  
FILING DATE: 09-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/629,602  
FILING DATE: 18-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/303001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154



```

QY .      122 IYALGPDGSGAGYEIMATDYTDVPASCLEKENEYAAGLPVRDYTT 166
          ||      |||  |  |  |  |  |  |  |  |  |  |  |  |  |
DB        91 -----QLMAQKITALNYSGL-NLNQYGVASPCRTICT 125

```

**RESULT 13**

hypothetical protein C37A5.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T19806  
R:White, S.  
submitted to the EMBL Data Library, March 1997  
A:Reference number: Z19180  
A:Accession: T19806  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-512 <TWL>  
A:Cross-references: EMBL:Z52828; PIDN:CA807337.1; GSPDB:GN00019; CESP:C37A5.1  
A:Experimental source: clone C37A5  
C:Genetics:  
A:Gene: CESP:C37A5.1  
A:Map position: 1  
A:Intons: 19/3, 60/2, 106/2, 162/1, 265/3, 313/3, 364/2, 388/1, 467/3  
C:Superfamily: *Caenorhabditis elegans* hypothetical protein C01H12.5

Query Match	8.68	Score	81.5	DB	2	Length	512
Best Local Similarity	26.50	Pred. NO.	15				
Matches	31	Conservative	17	Mismatches	46	Indels	23
						Gaps	5

```

OY 46 FTCTGGAAGCINLENDEKNVAMFMFMNNADTYQHFEEKATDCKMKGKNAEATYYQEDG 105
    111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 FTCTGTRHNDMEDTFTIDYFPRIL---TVQGFEE-----FAGMKV-----A 288
OY 106 QVLTDLVASFSDMCYIYALAPDGSAGYELMATPTTYTPVASCLEKFE--YANGLP 160
    111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 EALLNTPGEGDDDFECNLTIDRN-IATGAATYNSYTSYDPEMLADKFNDSIATYYP 344
    111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 14

S28179  
 alpha-amylase (EC 3.2.1.1) - Streptomyces lividans  
 C:Species: Streptomyces lividans  
 C:Date: 13-Jan-1995 #sequence\_rev1slon 13-Jan-1995 #text\_change 24-Apr-1998  
 C:Accession: S28179  
 R:Tseo, L.-S.; Lin, L.-L.; Chen, J.-C.; Hsu, W.H.  
 Biochim. Biophys. Acta 1171, 255-262, 1993  
 A:Title: Cloning and characterization of an alpha-amylase gene from Streptomyces lividans  
 A:Reference number: S28179; PMID:93144336; PMID:8424949  
 A:Accession: S28179  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-919 <TSA>  
 C:Genetics:  
 A:Start codon: GTG  
 C:Keywords: glycosidase; hydrolase

Query Match	8.6%	Score	81.5	DB	2	Length	919
Best Local Similarity	27.0%	Pred. No.	31				
Matches	34	Conservative	11	Mismatches	32	Indels	49
						Gaps	7

```

QY      23 KLVSENNDLTKA-----TYKNDPV-----WQNDTCV-----GTAAG 54
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      101 KGLTERKIDYKIGSTSIMMAPIFKNQPVQGTGRDASAGYHGWTIDETQVDPHGR 157
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY      55 NLNDEKEN-----VEAMFEMNNADTYQHTFEKAT-PDKMYGKNENATTYOD 104
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      158 --KKDKNLISRAHAGMKYF-----DYTNHTADYVDIEKSYDLSGAPPLYTKD 205
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY      105 GOVLTLD 110
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      210 GQPEDD 215
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

**RESULT 15**

hyp0538  
 Hypothetical protein MYPU\_2120 [Imported] - Mycoplasma pulmonis (strain UAB CTIP)  
 C:Species: Mycoplasma pulmonis  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
 C:Accession: D90538  
 R:Chenboud, I.; Hehlly, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, M.; Nucleic Acids Res 29, 2145-2153, 2001  
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma F  
 A:Reference number: A95312; MUID:21267165; PMID:11353064  
 A:Accession: D90538  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1377 <KUR>  
 A:Cross-references: GB:AI445566; PID:G14089625; PIDN:CACI3385.1; GSPDB:GN00153  
 A:Experimental source: strain UAB CTIP  
 C:Genetics:  
 A:Gene: MYPU\_2120  
 A:Genetic code: SGC3

Query Match	8.6%	Score 81.5	DB 2	Length 1377
Best Local Similarity	20.5%	Pred. No. 53		
Matches 41	Conservative 19	Mismatches 61	Indels 79	Gaps 8

```

OY 13 EHOQAMK-----HLOKLTVEENDL-----IKATYK 37
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1114 KRODAMKSEDKVYIDOKANOMKKTIASOYKLDQOKFKIRKNSKDALKETPLGTTINTLIN 1173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 38 NDPVWGNDFTCVGTAAQNLNDEDEKNEAFMFMMNADTVYOHTEEKATPDKMYGYNKEN- 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1174 EDKVMWADFA-----DKEIKDKIKRYK--DLEIQYSFE-----GYTNSNG 1212
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 97 ----AIYQTEGQVLTIVLAFSDNCYVITALGFDSGAGYELMATDYTDVPASCLEKF 152
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1213 EOIWWRNAQNEEAHVAHFKNLLSSDNNXKYGSLSVD-----QIWAFFKKRILVFAKFF 1266
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 153 NEYAAGLVPVRDVTSDCLPE 172
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1267 S-----NSDLEFPD 1274
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: August 1, 2003, 12:16:32  
Job time : 15.7771 secs

**THIS PAGE BLANK (USPTO)**



Db 288 AFTVLA-MMDYPTDFELGFLPANPVKGCRLLESEORITGLRALAGLVYNAAGSEHCY 346  
OY 121 VITAL-----GPDGSGAGYELMATDYTDVPASCLEKFNENYAGLPVRY 165  
Db 347 DIRLRLHSCADPTGCGTGPBARMDY---QACTE-INLTFASNNVTDMF 391

## RESULT 13

US-09-388-413B-6  
; Sequence 6, Application US/09388413B  
; Patent No. 6485955  
; GENERAL INFORMATION:  
; APPLICANT: Huber, Brigitte T.  
; TITLE OF INVENTION: QUIESCENT CELL DIPEPTIDYL PEPTIDASE: A NOVEL CYTOPLASMIC SERINE  
; FILE REFERENCE: 50420 CJP  
; CURRENT APPLICATION NUMBER: US/09/388,413B  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: 08/944,265  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-388-413B-6

Query Match 8.4%; Score 79.5; DB 4; Length 537;  
Best Local Similarity 25.9%; Pred. No. 3;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKLVEENYDLIKATYKNDPVGNDFTCVGTAAQNLNDEKNEVEMFMNN 72  
Db 221 REAFRQKDLFLGAGDVTYR-----W--EFGTC-----QPLS-DEKDLTQLFMRARN 264  
OY 73 ADTVYQHTFEKATPDKMYGYNENATYYQTE---DQGVLTDLVLA-----DDNCY 120  
Db 265 AFTVLA-MMDYPTDFELGFLPANPVKGCRLLESEORITGLRALAGLVYNAAGSEHCY 323  
OY 121 VITAL-----GPDGSGAGYELMATDYTDVPASCLEKFNENYAGLPVRY 165  
Db 324 DIRLRLHSCADPTGCGTGPBARMDY---QACTE-INLTFASNNVTDMF 368

## RESULT 14

US-08-284-941-2  
; Sequence 2, Application US/08284941  
; Patent No. 5863756  
; GENERAL INFORMATION:  
; APPLICANT: BARR, PHILIP J  
; APPLICANT: KIEFER, MICHAEL C  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND  
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM  
; STREET: FIVE PALO ALTO SQUARE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/284,941  
; FILING DATE: 2 August 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEELEY PH.D., RICHARD L.

; REGISTRATION NUMBER: 30092  
; REFERENCE/DOCKET NUMBER: CHIR-009/01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 843-5070  
; TELEFAX: (415) 857-0663  
; TELEFAX: 380816 COOLEY PA  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 969 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-284-941-2

Query Match 8.1%; Score 77.5; DB 2; Length 969;  
Best Local Similarity 21.3%; Pred. No. 12;  
Matches 39; Conservative 28; Mismatches 81; Indels 35; Gaps 7;

OY 13 EHODANKHLQKLVEENYDLIKATYKNDPVGNDFTCVGTAAQNLNDEKNEVEMFMNN 68  
Db 139 QOQEVKRVKRYQRSD---PQALYFNDPISNMWYLCGDKNRCRENNVQAAMKRGYT 195  
OY 69 FMNNDTVYQHTFEKATPDKMYGYNK-----ENATY-----QTEGCVLTDLVLA 113  
Db 196 GKNVYVYTIIDGIEENHFDLPAPNYDSYAVYDNGNDPSPRYDASNNENHGTCCAGEVA 255  
OY 114 FSDONCVIYALGPDGSGAGYELMATDYTDVPASCLEKFNENYAGLPVRY---DYVNSDC 169  
Db 256 ASANNSTCYIGIYNAKIGIRMLDGDVTYVE-----AKSLGIRPNYIDYISASW 306  
OY 170 LPE 172  
Db 307 GPD 309

## RESULT 15

US-08-447-642-2  
; Sequence 2, Application US/08447642  
; Patent No. 5989890  
; GENERAL INFORMATION:  
; APPLICANT: BARR, PHILIP J  
; APPLICANT: KIEFER, MICHAEL C  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND  
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM  
; STREET: FIVE PALO ALTO SQUARE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,642  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/284,941  
; FILING DATE: 2 August 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEELEY PH.D., RICHARD L.  
; REGISTRATION NUMBER: 30092  
; REFERENCE/DOCKET NUMBER: CHIR-009/01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 843-5070  
; TELEFAX: (415) 857-0663  
; TELEFAX: 380816 COOLEY PA  
; INFORMATION FOR SEQ ID NO: 2:

```

; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 518
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-518

Query Match
Best Local Similarity 25.9%; Pred. No. 2.7;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKLT-VEENYDLIKATYKNDPVWGNDF-TCVGTAAQNINDEKNEVAFMFPMNN 72
DB 219 REAFRQIKDLFLQAGAYDVR-----W-EFGTC-----QPLS-DEKDLTQLEMFARN 262
OY 73 ADYVQHTFEKATPDKMYGKKNATYQTE---DGOVLTDVLAFS-----DDNCY 120
DB 263 AFTVLA-MMDYPTDFELGPLPAPNVKVCRLISEAQRTIGLRALAGLVYNASGSEHCY 321
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFEVYAAGLPVRDYY 165
DB 322 DIYRLYHSCADPTGCGTGTPDARANDY---OACTE-INLTFASNNVTDMF 366

RESULT 10
US-09-794-236-2
; Sequence 2, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/794,236
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-236-2

Query Match
Best Local Similarity 8.4%; Score 79.5; DB 4; Length 492;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKLT-VEENYDLIKATYKNDPVWGNDF-TCVGTAAQNINDEKNEVAFMFPMNN 72
DB 221 REAFRQIKDLFLQAGAYDVR-----W-EFGTC-----QPLS-DEKDLTQLEMFARN 264
OY 73 ADYVQHTFEKATPDKMYGKKNATYQTE---DGOVLTDVLAFS-----DDNCY 120
DB 265 AFTVLA-MMDYPTDFELGPLPAPNVKVCRLISEAQRTIGLRALAGLVYNASGSEHCY 323
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFEVYAAGLPVRDYY 165
DB 324 DIYRLYHSCADPTGCGTGTPDARANDY---OACTE-INLTFASNNVTDMF 368

RESULT 11
US-09-345-469-1
; Sequence 1, Application US/09345469
; Patent No. 6369210
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
```

```

; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: 22012, A No. 6369210el Human Carboxypeptidase
; FILE REFERENCE: 5800-38
; CURRENT APPLICATION NUMBER: US/09/345,469
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-345-469-1

Query Match
Best Local Similarity 8.4%; Score 79.5; DB 4; Length 492;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKLT-VEENYDLIKATYKNDPVWGNDF-TCVGTAAQNINDEKNEVAFMFPMNN 72
DB 221 REAFRQIKDLFLQAGAYDVR-----W-EFGTC-----QPLS-DEKDLTQLEMFARN 264
OY 73 ADYVQHTFEKATPDKMYGKKNATYQTE---DGOVLTDVLAFS-----DDNCY 120
DB 265 AFTVLA-MMDYPTDFELGPLPAPNVKVCRLISEAQRTIGLRALAGLVYNASGSEHCY 323
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFEVYAAGLPVRDYY 165
DB 324 DIYRLYHSCADPTGCGTGTPDARANDY---OACTE-INLTFASNNVTDMF 368

RESULT 12
US-09-461-325-219
; Sequence 219, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 219
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-219

Query Match
Best Local Similarity 8.4%; Score 79.5; DB 4; Length 515;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKLT-VEENYDLIKATYKNDPVWGNDF-TCVGTAAQNINDEKNEVAFMFPMNN 72
DB 244 REAFRQIKDLFLQAGAYDVR-----W-EFGTC-----QPLS-DEKDLTQLEMFARN 287
OY 73 ADYVQHTFEKATPDKMYGKKNATYQTE---DGOVLTDVLAFS-----DDNCY 120
```

SEQUENCE CHARACTERISTICS:  
 LENGTH: 969 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-447-642-2

Query Match. 8.1%; Score 77.5; DB 2; Length 969;  
 Best Local Similarity 21.3%; Pred. No. 12; Mismatches 7;  
 Matches 39; Conservative 28; Indels 35; Gaps 7;

OY	13	EHODAMKHLQKLV	EEYNDLIKATYKNDP	VGNDFTCVGTAAONLNEDEKNVE-AW---FM	68
DB	139	QOQEVKRRVKKQVRSD---	PGALYFNDPIRNSNMWYLHCGDKNSRCRSEMANVQAAMKRGYT	195	
OY	69	FMNNADIVYQHTFEKATPDKMKGYNK----	ENAITY-----QTEGQVLTDLVA	113	
DB	196	GKNVYVYTLDDGIERNHDLAPNYDSYASYDVNGNDYDPSPRYDASNENKHGTRCAGHYA	255		
OY	114	FSDDMCYIYIALGPDGSGAGYEI	NATDVTDPASCLEKFN	EYAAGLPVR---DYYTSDC	169
DB	256	ASANSSTCIYIGIAYNAKIGIRMLDGDVTDVVE-----	AKSLGIRPNYIDLYSASW	306	
OY	170	LPE	172		
DB	307	GPD	309		

Search completed: August 1, 2003, 12:17:19  
 Job time: 16.1048 secs

**THIS PAGE BLANK (USPTO)**

ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-494  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1861 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-790-912-4

Query Match 8.68; Score 82; DB 2; Length 1861;  
Best Local Similarity 27.18; Pred. No. 10;  
Matches 51; Conservative 12; Mismatches 47; Indels 78; Gaps 12;

QY 7 DEANGHODAMKHLKLVEN-----YD-----LIKATYK 37  
DB 1370 DRRNNHR-AGELKKEFEDNAOETAKRORDHYDWRILDSGKEKLYRLNLYDAYK 1428  
QY 38 NDEWGNDFCTGVTAAQNLNEDENYEAEMFNNA-----DYVYQHTFEKATPD 87  
DB 1429 -----FDDDTY-----DKATVEAQFSSNPAMYKFFGPGVGNKKVYHNKGAVATGD 1474  
QY 88 KMY--GY---NKENATY---QTEDCQVLTDLVAFSDDCYVYALGPDG--SGAGYELW 137  
DB 1475 SVYKMGYRLDKDGAITYTHEMTHD-----SDNEYI---LEGYGRSGLGPEFR 1520  
QY 138 ATDYDVP 145  
DB 1521 AKGLQAP 1528

RESULT 3  
US-08-790-912-3  
Sequence 3, Application US/08790912  
Patent No. 5976542  
GENERAL INFORMATION:  
APPLICANT: Weiser, Jeffrey N.  
APPLICANT: Plaut, Andrew G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT  
TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-494  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1964 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-790-912-3

Query Match 8.58; Score 81; DB 2; Length 1964;  
Best Local Similarity 24.18; Pred. No. 14;  
Matches 45; Conservative 25; Mismatches 67; Indels 50; Gaps 8;

QY 12 GEHODAMKHLKLVEN-----YDLIKATYKNDPVWGNDFCTGTAQNLNEDENVEA 65  
DB 1073 GYMLDSQYVSKAEINPLTPVEPLSTSGKD-----SDFSKIATYQARALVYNNIEK 1128  
QY 66 WFMFMNADTYOHTFEKATPDKMYGNKENATYQTE-----DGOVLTDLVAFSD 117  
DB 1129 LLPFYN-----KSTYVYGNLVKENSLLYKELLSAVYMKMDQVITDIVS-NNQ 1176  
QY 118 NCYVYIALGPDGSGAGYEL-WATDYTDVPASCLEKNEVNAAG-----LPVRDY 165  
DB 1177 TANKLLHYNDHSEKFDLYQTFANLP-----EYNLGTGLYTPNDFLYDRDSI 1228  
QY 166 TSDCLPE 172  
DB 1229 VKRELP 1235

RESULT 4  
US-08-790-912-2  
Sequence 2, Application US/08790912  
Patent No. 5976542  
GENERAL INFORMATION:  
APPLICANT: Weiser, Jeffrey N.  
APPLICANT: Plaut, Andrew G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT  
TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:



```
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-989-545-14
```

```
Query Match      8.1%; Score 77; DB 10; Length 481;
Best Local Similarity 21.3%; Pred. No. 6.1;
Matches 40; Conservative 33; Mismatches 63; Indels 52; Gaps 9;
```

```
OY      2 KPWADDEAANGEHODANKHLQK--LVEENYDLIKATYKNDPVWGNDFTC-----VGTAA 53
Db      102 KPCQSDPEVPDGIKSASYKSEANNLIEE-----CEQAERLGAVD 141
OY      54 QNLNED-EKNVEAFMFEMNNADTVYQ-----HFEKATPDKNYG---YKNENA 97
Db      142 ESLSEETQKAVLQMTKHDSSDNFCEDADNICMWSKKKVFDIQSPAEYVDLLNPERY 201
OY      98 ITVQTEGQVLTVDVLAFSDDNCK---VIVALPGDGSAGYELWATDYTDVPAACLEK-- 151
Db      202 TGKKGPDAMKIMNVI--YEENCFRPOTIKRPLNPLASGCGTSEENTFTSWLEGICVYKRA 259
OY      152 FNEYAAGL 159
Db      260 FYRLISGL 267
```

```
Search completed: August 1, 2003, 12:23:44
Job time : 18.0362 secs
```





Db 263 AFTVLA-MMDYPTFLGRLPANPVKVCDDLSEAOQRTGRLALAGLVNNSGSEHCY 321  
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFNENYAGLPYRDVY 165  
Db 322 DLYRLHSCADPTGCGTGPDRAMDY---QACTE-INLTFASNVTDMF 366

RESULT 9  
US-10-068-134-1

; Sequence 1, Application US/10068134  
; Publication No. US20020156264A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Lidemann, Rosana  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Williamson, Mark  
; TITLE OF INVENTION: 22012, A NO. US20020156264A1el Human Carboxypeptidase  
; FILE REFERENCE: 5800-38  
; CURRENT APPLICATION NUMBER: US/10/068,134  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: US/09/345,469  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-068-134-1

Query Match  
Best Local Similarity 25.9%; Score 79.5; DB 14; Length 492;  
Pred. No. 3.3;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 QDAWKHLQKL-VEENYDLIKATYKNDPVWGNDP-TGCVTAQNLNDEKKNVEMFPMNN 72  
Db 221 REAFROIKDLFGAIDTVR-----W-EFGTC-----OPLS-DEKDLTQLFPMARN 264  
OY 73 ADVYQHTFEKATPDKMYGKKNENATTYOTE---DGOVLTDLAFS-----DDNXY 120  
Db 265 AFTVLA-MMDYPTFLGRLPANPVKVCDDLSEAOQRTGRLALAGLVNNSGSEHCY 323  
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFNENYAGLPYRDVY 165  
Db 324 DLYRLHSCADPTGCGTGPDRAMDY---QACTE-INLTFASNVTDMF 368

RESULT 10  
US-10-012-542-219

; Sequence 219, Application US/10012542  
; Publication No. US20030044851A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029P1  
; CURRENT APPLICATION NUMBER: US/10/012,542  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1999-12-14  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 60/089,507  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 1998-06-22  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 219  
; LENGTH: 515  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-012-542-219

Query Match  
Best Local Similarity 25.9%; Score 79.5; DB 15; Length 515;  
Pred. No. 3.6;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 QDAWKHLQKL-VEENYDLIKATYKNDPVWGNDP-TGCVTAQNLNDEKKNVEMFPMNN 72  
Db 244 REAFROIKDLFGAIDTVR-----W-EFGTC-----OPLS-DEKDLTQLFPMARN 287  
OY 73 ADVYQHTFEKATPDKMYGKKNENATTYOTE---DGOVLTDLAFS-----DDNXY 120  
Db 288 AFTVLA-MMDYPTFLGRLPANPVKVCDDLSEAOQRTGRLALAGLVNNSGSEHCY 346  
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFNENYAGLPYRDVY 165  
Db 347 DLYRLHSCADPTGCGTGPDRAMDY---QACTE-INLTFASNVTDMF 391

RESULT 11  
US-10-233-745-6

; Sequence 6, Application US/10233745  
; Publication No. US20030027282A1  
; GENERAL INFORMATION:  
; APPLICANT: Huber, Brigitte T.  
; TITLE OF INVENTION: QUIESCENT CELL DIPEPTIDYL PEPTIDASE: A NOVEL CYTOPLASMIC SERI  
; FILE REFERENCE: 50420 CIP  
; CURRENT APPLICATION NUMBER: US/10/233,745  
; CURRENT FILING DATE: 2002-09-03  
; PRIOR APPLICATION NUMBER: US/09/388,413  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: 08/944,265  
; PRIOR FILING DATE: 1997-10-06  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-233-745-6

Query Match  
Best Local Similarity 25.9%; Score 79.5; DB 15; Length 537;  
Pred. No. 3.8;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 QDAWKHLQKL-VEENYDLIKATYKNDPVWGNDP-TGCVTAQNLNDEKKNVEMFPMNN 72  
Db 221 REAFROIKDLFGAIDTVR-----W-EFGTC-----OPLS-DEKDLTQLFPMARN 264  
OY 73 ADVYQHTFEKATPDKMYGKKNENATTYOTE---DGOVLTDLAFS-----DDNXY 120  
Db 265 AFTVLA-MMDYPTFLGRLPANPVKVCDDLSEAOQRTGRLALAGLVNNSGSEHCY 323  
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFNENYAGLPYRDVY 165  
Db 324 DLYRLHSCADPTGCGTGPDRAMDY---QACTE-INLTFASNVTDMF 368

RESULT 12  
US-10-103-196-17

; Sequence 17, Application US/10103196  
; Publication No. US20030050466A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: TM4SF Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT008P1  
; CURRENT APPLICATION NUMBER: US/10/103,196  
; CURRENT FILING DATE: 2002-03-22



## RESULT 2

US-10-087-195-6

```
; Sequence 6, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGANISM: FS-HBP1
US-10-087-195-6
```

```
Query Match          100.0%; Score 952; DB 14; Length 172;
Best Local Similarity 100.0%; Pred. No. 4,4e-97;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DKPVMADEANGEHODAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNINDE 60
        |||
DB      1 DKPVMADEANGEHODAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNINDE 60

QY      61 KNEAFMFNNADTYOHTFEKATPDKNYGYNKRENAITYOTEDGVLTDVLAFSDN 120
        |||
DB      61 KNEAFMFNNADTYOHTFEKATPDKNYGYNKRENAITYOTEDGVLTDVLAFSDN 120

QY      121 VYIAGPDGSGAGYELMATDVTDPASCLEKFEFNEVAGLPVRDVTSDCLPE 172
        |||
DB      121 VYIAGPDGSGAGYELMATDVTDPASCLEKFEFNEVAGLPVRDVTSDCLPE 172
```

## RESULT 3

US-10-085-572-7

```
; Sequence 7, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 171
; TYPE: PRT
; ORGANISM: FS-HBP 2
US-10-085-572-7
```

```
Query Match          66.2%; Score 630; DB 14; Length 171;
Best Local Similarity 66.5%; Pred. No. 1,4e-61;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;
```

```
QY      1 DKPVMADEANGEHODAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNINDE 60
        |||
DB      1 NQDVMADDEANGAHQDAMKSLKADVENYVYKATKKNDPVWGNDFTCVGTAQNINDE 60
```

```
QY      61 KNEAFMFNNADTYOHTFEKATPDKNYGYNKRENAITYOTEDGVLTDVLAFSDN 120
        |||
DB      61 KSIQAEFLFMNNADTYOHTFEKATPDKNYGYNKRENAITYOTEDGVLTDVLAFSDN 120

QY      121 VYIAGPDGSGAGYELMATDVTDPASCLEKFEFNEVAGLPVRDVTSDCLPE 170
        |||
DB      121 VYIVPCTDNEBEGYELMATDVTDPASCLEKFEFNEVAGLPVRDVTSDCLPE 170
```

## RESULT 4

US-10-087-195-7

```
; Sequence 7, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 171
; TYPE: PRT
; ORGANISM: FS-HBP 2
US-10-087-195-7
```

```
Query Match          66.2%; Score 630; DB 14; Length 171;
Best Local Similarity 66.5%; Pred. No. 1,4e-61;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;
```

```
QY      1 DKPVMADEANGEHODAMKHLQKLVENYDLIKATYKNDPVWGNDFTCVGTAQNINDE 60
        |||
DB      1 NQDVMADDEANGAHQDAMKSLKADVENYVYKATKKNDPVWGNDFTCVGTAQNINDE 60

QY      61 KNEAFMFNNADTYOHTFEKATPDKNYGYNKRENAITYOTEDGVLTDVLAFSDN 120
        |||
DB      61 KSIQAEFLFMNNADTYOHTFEKATPDKNYGYNKRENAITYOTEDGVLTDVLAFSDN 120

QY      121 VYIAGPDGSGAGYELMATDVTDPASCLEKFEFNEVAGLPVRDVTSDCLPE 170
        |||
DB      121 VYIVPCTDNEBEGYELMATDVTDPASCLEKFEFNEVAGLPVRDVTSDCLPE 170
```

## RESULT 5

US-10-085-572-8

```
; Sequence 8, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 182
; TYPE: PRT
; ORGANISM: MS-HBP 1
US-10-085-572-8
```

```
Query Match          28.1%; Score 267.5; DB 14; Length 182;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:58 ; Search time 17.0362 seconds  
(without alignments)  
1199.017 Million cell updates/sec

Title: US-10-087-195-6  
Perfect score: 952  
Sequence: 1 DKPYMADEAANGEHQDAWKH.....NEXAAGLPVRDVTSDCLPE 172

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues  
Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	952	100.0	172	US-10-085-572-6	Sequence 6, Appl1
2	952	100.0	172	US-10-087-195-6	Sequence 6, Appl1
3	630	66.2	171	US-10-085-572-7	Sequence 7, Appl1
4	630	66.2	171	US-10-087-195-7	Sequence 7, Appl1
5	267.5	28.1	182	US-10-085-572-8	Sequence 8, Appl1
6	267.5	28.1	182	US-10-087-195-8	Sequence 8, Appl1
7	79.5	8.4	490	US-10-012-542-250	Sequence 250, App
8	79.5	8.4	490	US-10-012-542-518	Sequence 518, App
9	79.5	8.4	492	US-10-068-134-1	Sequence 1, Appl1
10	79.5	8.4	537	US-10-012-542-219	Sequence 219, App
11	79.5	8.4	537	US-10-233-745-6	Sequence 6, Appl1
12	77.5	8.1	268	US-10-103-196-17	Sequence 17, Appl1
13	77.5	8.1	969	US-09-961-403-6	Sequence 6, Appl1
14	77.5	8.1	1806	US-10-156-761-13509	Sequence 13509, A
15	77	8.1	481	US-09-989-545-14	Sequence 14, Appl1

16	76	8.0	221	9	US-09-728-914-4	Sequence 4, Appl1
17	76	8.0	1011	15	US-10-128-714-8129	Sequence 8129, Ap
18	75.5	7.9	468	9	US-09-726-643-51	Sequence 51, Appl
19	75.5	7.9	468	10	US-09-978-295A-337	Sequence 337, App
20	75.5	7.9	468	10	US-09-978-697-337	Sequence 337, App
21	75.5	7.9	468	10	US-09-978-192A-337	Sequence 337, App
22	75.5	7.9	468	10	US-09-999-832A-337	Sequence 337, App
23	75.5	7.9	468	11	US-09-978-189-337	Sequence 337, App
24	75.5	7.9	468	11	US-09-978-608A-337	Sequence 337, App
25	75.5	7.9	468	11	US-09-978-585A-337	Sequence 337, App
26	75.5	7.9	468	11	US-09-978-191A-337	Sequence 337, App
27	75.5	7.9	468	11	US-09-978-403A-337	Sequence 337, App
28	75.5	7.9	468	11	US-09-978-564A-337	Sequence 337, App
29	75.5	7.9	468	11	US-09-999-833A-337	Sequence 337, App
30	75.5	7.9	468	11	US-09-981-915A-337	Sequence 337, App
31	75.5	7.9	468	11	US-09-978-824-337	Sequence 337, App
32	75.5	7.9	468	11	US-09-918-985A-337	Sequence 337, App
33	75.5	7.9	468	11	US-09-978-423A-337	Sequence 337, App
34	75.5	7.9	468	11	US-09-978-193A-337	Sequence 337, App
35	75.5	7.9	468	11	US-09-999-830A-337	Sequence 337, App
36	75.5	7.9	468	11	US-09-978-757A-337	Sequence 337, App
37	75.5	7.9	468	11	US-09-978-187B-337	Sequence 337, App
38	75.5	7.9	468	11	US-09-978-643A-337	Sequence 337, App
39	75.5	7.9	468	12	US-09-978-375A-337	Sequence 337, App
40	75.5	7.9	468	14	US-10-042-141-51	Sequence 51, Appl
41	75.5	7.9	468	15	US-10-017-081A-337	Sequence 337, App
42	75.5	7.9	468	15	US-10-167-749-337	Sequence 337, App
43	75.5	7.9	468	15	US-10-013-921A-337	Sequence 337, App
44	75.5	7.9	468	15	US-10-013-929A-337	Sequence 337, App
45	75.5	7.9	468	15	US-10-016-177A-337	Sequence 337, App

ALIGNMENTS

RESULT 1  
US-10-085-572-6  
: Sequence 6, Application US/10085572  
: Publication No. US20020151499A1  
: GENERAL INFORMATION:  
: APPLICANT: Nuttall, Patricia, Ann  
: TITLE OF INVENTION: Treatment of Conjunctivitis  
: FILE REFERENCE: 2488-1-003  
: CURRENT APPLICATION NUMBER: US/10/085,572  
: PRIOR FILING DATE: 2002-02-27  
: PRIOR APPLICATION NUMBER: PCT/GB00/03282  
: PRIOR FILING DATE: 2000-08-24  
: PRIOR APPLICATION NUMBER: 9920674.0  
: NUMBER OF SEQ ID NOS: 8  
: SOFTWARE: FASTSEQ for Windows Version 4.0  
: SEQ ID NO: 6  
: LENGTH: 172  
: TYPE: PRT  
: ORGANISM: FS-HBPI  
US-10-085-572-6

Query Match	100.0%	Score 952	DB 14	Length 172
Best Local Similarity	100.0%	Pred. No. 4.4e-97		
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	DKPYMADEAANGEHQDAWKHKLIVERNYDLIKATYKNDPVGKNDFTCVGTAQNLNDEDE	60	
DB	1	DKPYMADEAANGEHQDAWKHKLIVERNYDLIKATYKNDPVGKNDFTCVGTAQNLNDEDE	60	
QY	61	KNYEAMFEMFNNDYQHTFEKATPDKMGYKKNENAITYQTEDGCVLTVLAFSDNCCY	120	
DB	61	KNYEAMFEMFNNDYQHTFEKATPDKMGYKKNENAITYQTEDGCVLTVLAFSDNCCY	120	
QY	121	VITALGPDGSGAGYEIMATDYTDVPASCLERFNEYAAGLPVRDVTSDCLPE	172	
DB	121	VITALGPDGSGAGYEIMATDYTDVPASCLERFNEYAAGLPVRDVTSDCLPE	172	

```

Query Match      9.3%; Score 88.5; DB 1; Length 919;
Best Local Similarity 20.1%; Pred. No. 4.1;
Matches 34; Conservative 30; Mismatches 64; Indels 41; Gaps 7;

OY 27 ENYDLIKATYNDPVMGNDFTCYGTAONLN-----EDEKNVEAMFMFMNND 74
Db 117 EDVSAVTKGMKSNPNPFDSPFNOSTHRTSFCSPDILPMFKDGASPNKVDW----- 168
OY 75 TVYQHTFEKAPRDKMKYGNKENATLYOTED---GQVLTDLVAF---SDDKCYAYALGP 127
Db 169 -MYDHTFEEDFQSSI-----ELDAVYFRLLDPLFNHSTNLSLIDVEIKRDGAVYFHLILGC 223

OY 128 DGSAGAGELMATDYTD-----VPASCLEKFNVEYAAGLVPVRDYTSD 168
Db 224 DTACHSYRPRYSAEYVDNWKYIDDOIPI-LIDKVNKFFADDTAAFIETFD 271

```

R.L. J. Bacteriol.: 183:3293-3302(2001).  
CC -I- FUNCTION: Oxidizes hypoxanthine and xanthine to uric acid.  
CC -I- CATALYTIC ACTIVITY: Xanthine + NAD(+)+ H2O = urate + NADH.  
CC -I- COFACTOR: Molybdopterin (by similarity).  
CC -I- PATHWAY: Purine catabolism.  
CC -I- SUBUNIT: Could be composed of four subunits: puccA, puccC, puccD and  
CC puceL.  
CC -I- INDUCTION: Expression is very low in excess nitrogen (glutamate  
CC plus ammonia) and is induced during limiting-nitrogen conditions  
CC (glutamate)' Expression decreases when allantoin is added during  
CC limiting-nitrogen conditions.  
CC -I- SIMILARITY: BELONGS TO THE XANTHINE DEHYDROGENASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 299120; CAB15238.1; -  
 DR PIR: B70017; B70017.  
 DR SUDLIS: B613989; PUCD.  
 DR InterPro: IPR000674; Aldxan\_ch\_C.  
 DR Pfam: PF02738; Ald\_Xan\_ch\_C2; 1.  
 DR Pfam: PF01315; Ald\_Xan\_ch\_C; 1.  
 KM Oxidoreductase; NAD; Molybdenum; Purine metabolism; Complete proteome.  
 SO SEQUENCE 745 AA; 80433 MW; CA3EC5F0094EFD CRC64;

Query Match 9.0%; Score 85.5; DB 1; Length 745;  
 Best Local Similarity 26.2%; Pred. No. 5.9;  
 Matches 32; Conservative 22; Mismatches 41; Indels 27; Gaps 6;

QY 22 QKVEENYDLIKATKNDPVGNDFTCVTAQNLNDEKNEVAMFMN-----A 73  
 DB 106 EEIAEALLEIOVEYKELEWDSPEKALRPNAQRLHEDGTLHRAF-FSNGDVEEGFQAS 164  
 QY 74 DTVYQHTFEKATPDKMGYNKEMAITVQTEGGVLTVDVLAFSODNCVYVIALGPDGSGAG 133  
 DB 165 DTVFEETYE--LPRQMHFY-----METEGVAVPE---DDGGFTMTA---GTQHG 206  
 QY 134 YE 135  
 DB 207 YK 208

RESULT 6  
 ID AMY\_STRLI STANDARD; PRT; 919 AA.  
 AC Q05884;  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DE Alpha-amylose precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE glucanohydrolase).  
 GN AMY.  
 OS Streptomyces lividans.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID:1916;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 34-43.  
 RC STRAIN-FK24;  
 RX MEDLINE-93144336; PubMed-8424949;  
 RA Tsao L.-S., Lin L.-L., Chen J.-C., Chen J.-H., Hsu W.-H.;  
 RT "Cloning and characterization of an alpha-amylose gene from  
 RT Streptomyces lividans.";  
 RL Biochim. Biophys. Acta 1171:255-262(1993).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE-93250042; PubMed-8485150;  
 RA Tsao L.-S., Lin L.-L., Chen J.-C., Chen J.-H., Hsu W.-H.;  
 RL Biochim. Biophys. Acta 1173:119-119(1993).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X70255; CAA49759.1; -

DR HSSP: P10529; 7TAA.  
 DR InterPro: IPR006589; Alp\_amy1\_cat\_sub.  
 DR InterPro: IPR006048; Alpha\_amy1\_C.  
 DR InterPro: IPR006047; Alpha\_amy1\_cat.  
 DR InterPro: IPR005323; PUD.  
 DR Pfam: PF00128; alpha-amylose; 1.  
 DR Pfam: PF02806; alpha-amylose\_C; 1.  
 DR Pfam: PF03714; PUD; 2.  
 DR SMART: SM00642; Amy; 1.  
 KM Hydrolyase; Glycosidase; Carbohydrate metabolism; Signal.  
 FT SIGNAL 1 33  
 FT CHAIN 34 919  
 FT ACT\_SITE 312 312  
 FT ACT\_SITE 316 316  
 FT ACT\_SITE 417 417  
 SO SEQUENCE 919 AA; 100642 MW; 035B59FB80F0BD10 CRC64;

Query Match 8.6%; Score 81.5; DB 1; Length 919;  
 Best Local Similarity 27.0%; Pred. No. 17;  
 Matches 34; Conservative 11; Mismatches 32; Indels 49; Gaps 7;

QY 23 KIVEENYDLIKA-----TKNDPY-----WGNDFCV-----GTAQ 54  
 DB 101 KGLTEKLDYIKGJSTISMAPIFKNPVGGTGKASAGYHGWTIDFTVDPHFGT--- 157  
 QY 55 NLNDEKN-----VEAMFMNNDATVYQHTFEKAT-PDKMGYNKEMAITVQTE 104  
 DB 158 --NKDLKNLISKAKMKVVF-----DVTNHTADVVDYEEKSYLSKGFPIYTKD 209  
 QY 105 GQVLTG 110  
 DB 210 GQPFDD 215

RESULT 7  
 ID HTRE\_ECOLI STANDARD; PRT; 865 AA.  
 AC P33129;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Outer membrane usher protein htre precursor (Heat shock protein E).  
 GN HTRE OR B0139.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID:562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE-93352405; PubMed-8102362;  
 RA Raina S., Missiakas D., Baird L., Kumar S., Georgopoulos C.;  
 RT "Identification and transcriptional analysis of the Escherichia coli  
 RT htre operon which is homologous to pap and related pilin operons.";  
 RL J. Bacteriol. 175:5009-5021(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / W3110;  
 RX MEDLINE-94261430; PubMed-8202364;  
 RA Fujita N., Mori H., Yura T., Ishihama A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT the 2.4-4.1 mb (110,917-193,643 bp) region.";  
 RL Nucleic Acids Res. 22:1637-1639(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE-97426517; PubMed-9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Borna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).

```

RN [4]
RP SEQUENCE OF 30-33.
RA Raina S.:
RL Submitted (DEC-1994) to the SWISS-PROT data bank.
CC -1- FUNCTION: PROBABLE PORIN-LIKE PROTEIN NECESSARY FOR THE ASSEMBLY
CC OF A PILIN-TYPE PROTEIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
CC (By similarity).
CC -1- INDUCTION: By heat shock.
CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 481 WHICH CREATES TWO SEPARATE ORFS.
CC -1- CAUTION: REF.1 SEQUENCE WAS INCORRECT IN POSITION 861 ONWARD DUE
CC TO CLONING ARTIFACT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L00680; AAA23721.1; ALT-SEQ.
CC EMBL: D26562; -; NOT_ANNOTATED_CDS.
CC DR EMBL: AE000123; AAC73250.1; -.
CC DR PIR: C64737; C64737.
CC DR EcGene: EG11972; htre.
CC DR InterPro: IPR000015; Fimb_usher.
CC DR Pfam: PF00577; Usher; 1.
CC DR PROSITE: PS01151; FIMBRIAL-USHER; 1.
CC KW Outer membrane; Transmembrane; Fimbria; Transport; signal; Heat shock;
CC Complete proteome.
CC FT SIGNAL 1 29
CC FT CHAIN 30 865 OUTER MEMBRANE USHER PROTEIN HTRE.
CC FT DISULFID 838 862
CC FT CONFLICT 51 51 S -> T (IN REF. 1).
CC FT CONFLICT 810 811 OG -> HR (IN REF. 1).
CC FT CONFLICT 849 849 E -> P (IN REF. 1).
CC SQ SEQUENCE .865 AA; 95499 MW; F1748B551E45AAE CRC64;
CC -----
Query Match 8.58; Score 80.5; DB 1; Length 865;
Best Local Similarity 19.38; Pred. No. 19;
Matches 46; Conservative 33; Mismatches 60; Indels 99; Gaps 12;
QY 5 WADEAENGEDDAWKHLQKLVENYDLIKATYKNDPWGNDFTCVGTAQNLNDEKNVE 64
DB 530 WSDVWASGQNR-----SNYSL---GYNSTSGSYV---SAQRSMNEDGDTDD 572
QY 65 AMFW-----FMNNADTVYQ-----78
DB 573 SVYLSFTPIETKLGTEORTSGFOSIDTQISSDFKNNQNLNVSSSGYSDNARVSYVNTG 632
QY 79 HTEFKATPPDKMY--GY-----AKENA--ITYOTEDGOVL--TDVLASDD 117
DB 633 YTMNKAAGKDSLGVGVGYASYSPWGTLAGSISANDNSRQVSLSTDGFFVLHSGGLTFSND 692
QY 118 N-----CYVYALPGDGSAGY-----ELWATDYTPVPAACLEKFNFEYAGLPVPRV 164
DB 693 SFSQSDTLAVYQAFAGAGARINYNSTIDRWGIVT-----SALSPYHENRALALINDI 746

```

```

OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 14028;
RX MEDLINE-95394809; Pubmed-7655482;
RA Gunn J.S., Alpuche-Aranda C.M., Loomis W.P., Belden W.J., Miller S.I.;
RT "Characterization of the salmonella typhimurium pagC/pagD chromosomal
RT region.";
RL J. Bacteriol. 177:5040-5047(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LT2 / SSGC1412 / ATCC 700720;
RX MEDLINE-21534948; Pubmed-11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Potwolk S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Gwral N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U31849; AA82997.1; -.
CC DR EMBL: AE008754; AAL20169.1; -.
CC DR StyGene: SG10627; envF.
CC DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC KW Membrane; Lipoprotein; Signal; Complete proteome.
CC FT SIGNAL 1 25
CC FT CHAIN 26 262 PROBABLE LIPOPROTEIN ENVE.
CC FT LIPID 26 26 N-ACYL DIGLYCERIDE (POTENTIAL).
CC FT CONFLICT 223 223 D -> A (IN REF. 1).
CC FT CONFLICT 245 262 QRRKAAKMNRIQDTEK -> PA (IN REF. 1).
CC SQ SEQUENCE 262 AA; 29572 MW; ZBC6042085839143 CRC64;
CC -----
Query Match 8.48; Score 80; DB 1; Length 262;
Best Local Similarity 23.78; Pred. No. 5.1;
Matches 27; Conservative 21; Mismatches 46; Indels 20; Gaps 5;
QY 38 NDPVWGN---DFTCVGTAQNLNDEKNVEAMFPMNNADTV-----YQHTFEKAT 85
DB 110 NNGVWGQSPVNLITFTITREE-QKEALKALKF---DSINFSITPERIOETIRAN 164
QY 86 PDKMYGNKENVATYQTE--DGOVLVDYALFSDNCEYVYALPGDGSAGYEL 136
DB 165 ASGIISVTEDSDIVRAEIAHNGEFYDITTAARTARAATLTKDGSIAGYEI 218

```

RESULT 8  
ID ENVF\_SALTY STANDARD: PRT; 262 AA.  
AC Q56032;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable lipoprotein envf precursor.  
GN ENVF OR STM1240.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

RESULT 9  
ID ARJ2\_ANASP STANDARD: PRT; 387 AA.  
AC Q8YF9;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Arginine biosynthesis bifunctional protein argJ 2 [Includes: Glutamate  
N-acetyltransferase (EC 2.3.1.35) (Ornithine acetyltransferase)  
(Ornithine transacetylase) (OATase); Amino-acid acetyltransferase  
(EC 2.3.1.1) (N-acetylglutamate synthase) (AGS)] [Contains: Arginine  
biosynthesis bifunctional protein argJ2 alpha chain; Arginine  
biosynthesis bifunctional protein argJ2 beta chain].  
GN ARGJ2 OR ALR4235.  
OS Anabaena sp. (strain PCC 7120).



CC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Molk G.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimp S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium *Anabaena* sp. strain FCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 CC -1- FUNCTION: Catalyzes two activities which are involved in the  
 CC cyclic version of arginine biosynthesis: the synthesis of  
 CC acetylglutamate from glutamate and acetyl-CoA, and of ornithine by  
 CC transacylation between acetylornithine and glutamate (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate = L-  
 CC ornithine + N-acetyl-L-glutamate.  
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + L-glutamate = CoA + N-acetyl-L-  
 CC glutamate.  
 CC -1- PATHWAY: Arginine biosynthesis; first step.  
 CC -1- SUBUNIT: Heterotrimer of two alpha and two beta chains (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- MISCELLANEOUS: Some bacteria possess a monofunctional argu, i.e.,  
 CC capable of catalyzing only the fifth step of the arginine  
 CC biosynthetic pathway.  
 CC -1- SIMILARITY: Belongs to the argu family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP003595; BAB75934.1; -  
 DR PIR: AD2335; AD2335.  
 DR HAMAP: MF\_01106; -; 1.  
 DR InterPro: IPR002813; ArgJ.  
 DR Pfam: PF01960; ArgJ; 1.  
 DR ProDom: PD004193; ArgJ; 1.  
 DR TIGRPFAMs: TIGR00120; ArgJ; 1.  
 DR Arginine biosynthesis; Multifunctional enzyme; Transferase;  
 KW Acytransferase; Complete proteome.  
 FT CHAIN 1 179  
 FT ARGININE BIOSYNTHESIS BIFUNCTIONAL  
 FT PROTEIN ARGJ2 ALPHA CHAIN (BY  
 FT SIMILARITY).  
 FT ARGININE BIOSYNTHESIS BIFUNCTIONAL  
 FT PROTEIN ARGJ2 BETA CHAIN (BY  
 FT SIMILARITY).  
 FT CHAIN 180 387  
 FT ARGININE BIOSYNTHESIS BIFUNCTIONAL  
 FT PROTEIN ARGJ2 BETA CHAIN (BY  
 FT SIMILARITY).  
 FT SITE 179 180 CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).  
 FT SEQUENCE 387 AA; 41145 MW; B4A340DC3FF3013D CRC64;  
 SO  
 Query Match 8.4%; Score 80; DB 1; Length 387;  
 Best Local Similarity 25.5%; Pred. No. 8.2;  
 Matches 36; Conservative 23; Mismatches 50; Indels 32; Gaps 8;  
 OY 4 YMADEAANGHODAMKHLQKLVENYDLK-ATYKNDPWPNGDFTCVGTAONLNDENK 62  
 DB 267 VYDSDAAN-YAQAHRVAKAIV--NSPLVKTAVYAGADPMWGRVAMAIGKC-----BDEERD 317  
 OY 63 VEAFMFEMNADIVYCHTEFEKATPPDKMGYKNKNATLYGTEDQCVLTDLVAFSDDCYVI 122  
 DB 318 I-----NPDQVVRP-----DEVQYIPN---TFQENDEKLEKTSKEKVINIHS 359  
 OY 123 YALGPDGSGAGYELMATDYTD 143  
 DB 360 LNIQTDVA-----TWGCDLLE 376

RESULT 10  
 SYSCT\_YEAST  
 ID SYSCT\_YEAST STANDARD; PRT; 462 AA.  
 AC P07284;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Seru1-trna synthetase, cytoplasmic (EC 6.1.1.11) (Serine--trna  
 DE ligase) (SeruS).  
 GN SER1 OR SER5 OR YDR023W OR YD9813.01.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87174725; PubMed=3031581;  
 RA Weygand-Burassevic I., Johnson-Burke D., Soell D.;  
 RT "Cloning and characterization of the gene coding for cytoplasmic  
 RT seru1-trna synthetase from *Saccharomyces cerevisiae*.";  
 RL Nucleic Acids Res. 15:1887-1904(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Bowman S., Barrell B.G., Rajandream M.A.;  
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051598; PubMed=8896275;  
 RA Elde L.G., Sander C., Prydz H.;  
 RT "Sequencing and analysis of a 35.4 kb region on the right arm of  
 RT chromosome IV from *Saccharomyces cerevisiae* reveal 23 open reading  
 RT frames.";  
 RL Yeast 12:1085-1090(1996).  
 RN [4]  
 RP SEQUENCE OF 451-462 FROM N.A.  
 RC STRAIN=S288c;  
 RA Polley L.S., Fox T.D.;  
 RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-serine + trna(Ser) = AMP + diphosphate  
 CC + L-seru1-trna(Ser).  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-trna synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04884; CA28572.1; -  
 DR EMBL: Z47814; CA67802.1; -  
 DR EMBL: X95966; CA65216.1; -  
 DR EMBL: Z74319; CA38844.1; -  
 DR EMBL: L15408; AAC37412.1; -  
 DR PIR: S50930; YSBYC.  
 DR HSSP: P34945; ISER.  
 DR SGD: S0002430; SER1.  
 DR GO: GO:0005737; C:cytoplasm; IDA.  
 DR GO: GO:0004828; F:serine-trna ligase activity; IDA.  
 DR InterPro: IPR002314; trna-synt\_2b.  
 DR InterPro: IPR002317; trna-synt\_ser.  
 DR InterPro: IPR006195; trna\_ligase\_II.  
 DR Pfam: PF02403; Seru1-trna\_N; 1.  
 DR Pfam: PF00587; trna-synt\_2b; 1.  
 DR PRINTS: PR00981; TRNASYNTHSER.  
 DR TIGRPFAMs: TIGR00414; serS; 1.  
 DR PROSITE: PS50862; AA\_TRNA\_LIGASE\_II; 1.



RT	FT	AMINOACID	SEQUENCE	SEQUENCE FROM N.A.	AND PARTIAL SEQUENCE.
RT	FT	CONFLICT	224	224	L -> P (LN REF 1)
SO	SEQUENCE	462 AA	53309 MM	BBD6DEA476F547B	CRC664
Query Match		8.4%	Score 80	DB 1	Length 462
Best Local Similarity		24.6%	Pred. No. 10		
Matches	33	Conservative	17	Mismatches	40
				Indels	44
				Gaps	7
Oy		13	EHQDAWKHLQALV	---EENYDIKATYKNDPVGNDFTCVGRAQQLNE	---DEKNVEAM 66
Db		308	EPKMSWEFEKEMISYSEEFKSLKLPYR	-----IVGIVSGELNNAAKKYLDEAM	357
Oy		67	FMFM-----NNADTVYQHTFE	-----KATPDKMYGY	-----NKENATIT 99
Db		358	FPLYQEKYKELVSCSNCTDYQSNLNLEIRCSIKMGDBREKKYVHCINSLTATQALACCILE		417
Oy		100	-YQTEGQVLDLV 112		
Db		418	NYQTEDEGLVPEVL 431		
RESULT 11					
DP22_HUMAN					
ID	DP22_HUMAN	STANDARD	PRT	492 AA	
AC	Q9UHL4	Q969X4			
DT	16-OCT-2001	(rel. 40, Created)			
DT	28-FEB-2003	(rel. 41, Last sequence update)			
DE	15-SEP-2003	(rel. 42, Last annotation update)			
DE	Dipeptidyl-peptidase II precursor (EC 3.4.14.2) (DPP II) (Dipeptidyl				
DE	aminopeptidase II) (Quiescent cell proline dipeptidase) (Dipeptidyl				
DE	peptidase 7)				
GN	DP22 OR DP27 OR OPP.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_Taxid=9606				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RP	MEDLINE=20036542; PubMed=10567372;				
RX	Underwood R., Chitravuri M., Lee H., Schmitz T., Kabcenell A.K.,				
RT	Yardley Y., Huber B.T.;				
RT	*Sequence, purification, and cloning of an intracellular serine				
RL	protease, quiescent cell proline dipeptidase.*;				
RL	J. Biol. Chem. 274:34053-34058(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RP	TISSUE=Brain;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,				
RA	Altshul S.F., Zeeberg B., Buecaw K.H., Scheifer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Ditachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,				
RA	Brownstein M., Usdin T.B., Toshiyuki S., Carninci P., Prance C.,				
RA	Raba S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullany S.J.,				
RA	Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,				
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakeley R.C., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield V.S.N., Krzywinski M.T., Skalska U., Smalhus D.E.,				
RA	Schmerer A., Schein J.E., Jones S.J.M., Marra M.A.;				
RT	*Generation and initial analysis of more than 15,000 full-length				
RT	human and mouse cDNA sequences.*;				
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[3]				
RP	CHARACTERIZATION.				
RX	MEDLINE=20521747; PubMed=11067927;				
RA	Chitravuri M., Aggaraberes F., Mathieu S.L., Lee H., Huber B.T.;				
RT	*Vesicular localization and characterization of a novel				

Query Match	Best Local Similarity	8.4%	Score 79.5:	DB 1:	Length 492:				
Matches	44:	Conservative	27:	Mismatches	50:	Indels	41:	Gaps	12:
Qy	15	ODAMKHLCKL-VEENYDILKATYKNDPVWGDNF-TCVTAONAONLMEKXNEAMFPMNN	72						
Db	221	REARFOITDLQGLQYDVR-----W--EFGTC-----QPLS-DKDKLTQLFMRNR	264						
Qy	73	ADVYQHTFEKATPDKMYKGNATYQTE---DGGVLDVLAFS-----DDNCY	120						
Db	265	AFYVLA-IMDYPYDFDLGPIRANPVKRGKCRLLSEAGRTIGRLAGLAVYNASGSEHCY	323						
Qy	121	VIVYL-----GPDGAGYELMARDYTPVAPASCLEKFMFVNAAGLPVRQVY	165						
Db	324	DIYRLYHSCADPTGGGTGPDARAWY---OACTE-INLTPRSNNVYDMF	368						

DE Phosphatidylinositol 3-kinase tor1 (EC 2.7.1.137) (PI3-kinase)  
 (Ptdins-3-kinase) (PI3K).  
 GN TOR1 OR SPC30D10.10C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 CC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Spouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brocks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Haldago J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellon J., Simmonds M., Squares R., Stevens K., Sharp S.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voickert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Wellens I., Vansteede E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.K., Paulsen I., Potashkin J.,  
 RA Siparovsky G.V., Ussery D., Barrell B.G., Nurse P.;  
 RA "The genome sequence of Schizosaccharomyces pombe.";  
 RT Nature 415:871-880(2002).  
 RN [2]  
 RP IDENTIFICATION, FUNCTION, AND INDUCTION.  
 RX MEDLINE=21269225; PubMed=11096119;  
 RA Weisman M., Choder M.;  
 RT "The fission yeast TOR homolog, tor1+, is required for the response to  
 RT starvation and other stresses via a conserved serine.";  
 RT J. Biol. Chem. 276:7027-7032(2001).  
 CC -1- FUNCTION: phosphatidylinositol 3-kinase homolog required for G1  
 CC progression and entry into stationary phase. Also required for the  
 CC onset of meiosis and sporulation under nitrogen and carbon  
 CC starvation conditions.  
 CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +  
 CC 1-phosphatidyl-1D-myo-inositol 3-phosphate.  
 CC -1- INDUCTION: By nitrogen and/or carbon starvation, cold, osmotic and  
 CC oxidative stress.  
 CC -1- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.  
 CC -1- SIMILARITY: Contains 16 HEAT repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z97992; CAB10805.1; -;  
 DR PIR: T40186; T40186.  
 DR HSSP: P42345; IFAP.  
 DR GeneDB SPombe: SPC30D10.10c; -;  
 DR InterPro: IPR003151; FAT.  
 DR InterPro: IPR003152; FATC.  
 DR InterPro: IPR000357; HEAT\_repeat.  
 DR InterPro: IPR000403; PI3\_PI4\_kinase.  
 DR Pfam: PF02259; FAT; 1.

DR Pfam: PF02260; FATC; 1.  
 DR Pfam: PF00454; PI3\_PI4\_kinase; 1.  
 DR SMART: SM00146; PI3K; 1.  
 DR PROSITE: PS50077; HEAT\_REPEAT; FALSE\_NEG.  
 DR PROSITE: PS00915; PI3\_4\_KINASE\_1; 1.  
 DR PROSITE: PS00916; PI3\_4\_KINASE\_2; 1.  
 DR PROSITE: PS50290; PI3\_4\_KINASE\_3; 1.  
 KW Transferase; Kinase; Cell cycle; Repeat.  
 FT REPEAT 1  
 FT REPEAT 164 201 HEAT 1.  
 FT REPEAT 331 371 HEAT 2.  
 FT REPEAT 410 449 HEAT 3.  
 FT REPEAT 474 512 HEAT 4.  
 FT REPEAT 522 560 HEAT 5.  
 FT REPEAT 562 596 HEAT 6.  
 FT REPEAT 642 679 HEAT 7.  
 FT REPEAT 684 722 HEAT 8.  
 FT REPEAT 728 766 HEAT 9.  
 FT REPEAT 843 880 HEAT 10.  
 FT REPEAT 904 923 HEAT 11.  
 FT REPEAT 924 961 HEAT 12.  
 FT REPEAT 964 1003 HEAT 13.  
 FT REPEAT 1005 1042 HEAT 14.  
 FT REPEAT 1732 1769 HEAT 15.  
 FT REPEAT 1987 2335 HEAT 16.  
 SQ SEQUENCE 2335 AA; 266181 MW; 5DCFFCF4B8B9A4 CRC64;  
 Query Match 8.48; Score 79.5; DB 1; Length 2335;  
 Best Local Similarity 22.68; Pred. No. 78;  
 Matches 43; Conservative 29; Mismatches 55; Indels 63; Gaps 11;  
 QY 12 GENDADKHKQK---LVEENYDLIRTKYNDYWGNDPFCVGTAAQNLNDE-----K 61  
 DB 1402 GQYKKAIGHIERRDILVNLSSITGESYR--AYG-----INVSQMSLELEITIDYK 1454  
 QY 62 NVAMWFNNADVTYQHTFEKATPDKMYGKNE-----NATYQTEGQ-----V 107  
 DB 1455 NMQ---YENNIDSL-----KTMWRKLECCQKNVDVWNTLFRALVYSPDSEPMIK 1505  
 QY 108 LTVLAESD---DNCYVIYALPDGSGAGE-----LMAIDYDVP 145  
 DB 1506 LADICRSRSDIKLSNOCITLMGRDPSNA--YPLDSLKLNPVVYTYLKYLMATDQKINIA 1564  
 QY 146 ASCLKEFNEY 155  
 DB 1565 VSELEEFYST 1574  
 RESULT 13  
 PFTA\_SCHPO  
 ID PFTA\_SCHPO STANDARD; PRT; 294 AA.  
 AC 060052;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Protein farnesyltransferase alpha subunit (EC 2.5.1.-) (CAAX  
 DE farnesyltransferase alpha subunit) (RAS proteins prenyltransferase  
 DE alpha) (Frase-alpha).  
 GN CWP1 OR SPAPB1A10.04C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 CC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBUNIT.  
 RC STRAIN=972;  
 RX MEDLINE=98453135; PubMed=97818174;  
 RA Atelliano M., Coll P.M., Yang W., Duran A., Tamanoi F., Perez P.;  
 RT "Characterization of the geranylgeranyl transferase type I from  
 RT Schizosaccharomyces pombe.";  
 RT Mol. Microbiol. 29:1357-1367(1998).  
 RN [2]



RX MEDLINE-97064242; PubMed-8906861;  
RA Zhong M., Benjannet S., Lazure C., Munzer S., Seldah N.G.;  
RT "Functional analysis of human PACE4-A and PACE4-C isoforms:  
RT Identification of a new PACE4-CS isoform.";  
RL FEBS Lett. 396:31-36(1996).  
RN [8]  
RP CHARACTERIZATION.  
RX MEDLINE-99233559; PubMed-10215603;  
RA Susic J.F., Moehring J.M., Innocencio N.M., Lucchini J.W.,  
RA Moehring T.J.;  
RT "Endoprotease PACE4 is Ca<sup>2+</sup>-dependent and temperature-sensitive and  
RT can partly rescue the phenotype of a furin-deficient cell strain.";  
RL Blochem. J. 339:639-647(1999).  
RN [9]  
RP PROCESSING.  
RX MEDLINE-98408849; PubMed-9738469;  
RA Nagahama M., Taniguchi T., Hashimoto E., Imamaki A., Mori K.,  
RA Tsuji A., Matsuda Y.;  
RT "Biosynthetic processing and quaternary interactions of proprotein  
RT convertase SPC4 (PACE4).";  
RL FEBS Lett. 434:155-159(1998).  
CC -1- FUNCTION: LIKELY TO REPRESENT AN ENDOPROTEASE ACTIVITY WITHIN THE  
CC CONSTITUTIVE SECRETORY PATHWAY, WITH UNIQUE RESTRICTED  
CC DISTRIBUTION IN BOTH NEUROENDOCRINE AND NON-NEUROENDOCRINE TISSUES  
CC AND CAPABLE OF CLEAVAGE AT THE RX(R/R) CONSENSUS MOTIF.  
CC -1- CATALYTIC ACTIVITY: Release of mature proteins from their  
CC proproteins by cleavage of Arg-Xaa-Yaa-Aro-I'-Zaa bonds, where Xaa  
CC can be any amino acid and yaa is Arg or Lys.  
CC -1- COFACTOR: PACE4A IS PROBABLY CALCIUM-DEPENDENT.  
CC -1- SUBUNIT: THE PACE4A-I PRECURSOR PROTEIN SEEMS TO EXIST IN THE  
CC RETICULUM ENDOPLASMIC AS BOTH A MONOMER AND A DIMER-SIZED COMPLEX  
CC WHEREAS MATURE PACE4A-I EXISTS ONLY AS A MONOMER, SUGGESTING THAT  
CC PROPEPTIDE CLEAVAGE AFFECTS ITS TERTIARY OR QUATERNARY STRUCTURE.  
CC -1- SUBCELLULAR LOCATION: PACE4A-I AND PACE4-AII ARE SECRETED. PACE4C  
CC AND PACE4CS ARE NOT SECRETED AND REMAIN RETAINED IN ZYMOGENIC FORM  
CC IN ENDOPLASMIC RETICULUM. PACE4E-I AND PACE4E-II ARE RETAINED  
CC INTRACELLULARLY PROBABLY THROUGH A HYDROPHOBIC CLUSTER IN THEIR C-  
CC TERMINUS. PACE4B MIGHT BE SECRETED.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named Isoforms=8;  
CC Name-PACE4A-I; Synonyms-PACE4;  
CC IsoId-P29122-1; Sequence-Displayed;  
CC Name-PACE4A-II; Sequence-Sequence=VSP\_005436;  
CC IsoId-P29122-2; Sequence-Sequence=VSP\_005436;  
CC Name-PACE4B; Synonyms-PACE4.1;  
CC IsoId-P29122-3; Sequence-Sequence=VSP\_005428, VSP\_005429;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4C;  
CC IsoId-P29122-4; Sequence-Sequence=VSP\_005432, VSP\_005433;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4CS;  
CC IsoId-P29122-5; Sequence-Sequence=VSP\_005430, VSP\_005431;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4D;  
CC IsoId-P29122-6; Sequence-Sequence=VSP\_005427, VSP\_005434, VSP\_005435;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4E-I;  
CC IsoId-P29122-7; Sequence-Sequence=VSP\_005437;  
CC Name-PACE4E-II;  
CC IsoId-P29122-8; Sequence=VSP\_005436, VSP\_005437;  
CC -1- TISSUE SPECIFICITY: EACH PACE4 ISOBFORM EXHIBITS A UNIQUE  
CC RESTRICTED DISTRIBUTION. PACE4A-I IS EXPRESSED IN HEART, BRAIN,  
CC PLACENTA, LUNG, SKELETAL MUSCLE, KIDNEY, PANCREAS, BUT AT  
CC COMPARATIVELY HIGHER LEVELS IN THE LIVER. PACE4A-II IS AT LEAST  
CC EXPRESSED IN PLACENTA. PACE4B WAS ONLY FOUND IN THE EMBRYONIC  
CC KIDNEY CELL LINE FROM WHICH IT WAS ISOLATED. PACE4C AND PACE4D ARE  
CC EXPRESSED IN PLACENTA. PACE4E-I IS EXPRESSED IN CEREBELLUM,  
CC PLACENTA AND PITUITARY. PACE4E-II IS AT LEAST PRESENT IN  
CC CEREBELLUM.  
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE  
CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC  
CC RETICULUM. ISOBFORM PACE4D LACKS THE PROPEPTIDE DOMAIN.  
CC -1- SIMILARITY: Belongs to peptidase family 58.

CC -1- SIMILARITY: Contains 1 homo B/P domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M80482; AAA5998.1; -  
DR EMBL; AB001914; BAA21620.1; -  
DR EMBL; AB001898; BAA21620.1; JOINED.  
DR EMBL; AB001901; BAA21620.1; JOINED.  
DR EMBL; AB001903; BAA21620.1; JOINED.  
DR EMBL; AB001902; BAA21620.1; JOINED.  
DR EMBL; AB001904; BAA21620.1; JOINED.  
DR EMBL; AB001905; BAA21620.1; JOINED.  
DR EMBL; AB001906; BAA21621.1; JOINED.  
DR EMBL; AB001907; BAA21621.1; JOINED.  
DR EMBL; AB001908; BAA21621.1; JOINED.  
DR EMBL; AB001909; BAA21621.1; JOINED.  
DR EMBL; AB001910; BAA21621.1; JOINED.  
DR EMBL; AB001911; BAA21621.1; JOINED.  
DR EMBL; AB001912; BAA21621.1; JOINED.  
DR EMBL; AB001913; BAA21621.1; JOINED.  
DR EMBL; AB001914; BAA21621.1; JOINED.  
DR EMBL; AB001915; BAA21621.1; JOINED.  
DR EMBL; AB001916; BAA21621.1; JOINED.  
DR EMBL; AB001917; BAA21621.1; JOINED.  
DR EMBL; AB001918; BAA21621.1; JOINED.  
DR EMBL; AB001919; BAA21621.1; JOINED.  
DR EMBL; AB001920; BAA21621.1; JOINED.  
DR EMBL; AB001921; BAA21621.1; JOINED.  
DR EMBL; AB001922; BAA21621.1; JOINED.  
DR EMBL; AB001923; BAA21621.1; JOINED.  
DR EMBL; AB001924; BAA21621.1; JOINED.  
DR EMBL; AB001925; BAA21621.1; JOINED.  
DR EMBL; AB001926; BAA21621.1; JOINED.  
DR EMBL; AB001927; BAA21621.1; JOINED.  
DR EMBL; AB001928; BAA21621.1; JOINED.  
DR EMBL; AB001929; BAA21621.1; JOINED.  
DR EMBL; AB001930; BAA21621.1; JOINED.  
DR EMBL; AB001931; BAA21621.1; JOINED.  
DR EMBL; AB001932; BAA21621.1; JOINED.  
DR EMBL; AB001933; BAA21621.1; JOINED.  
DR EMBL; AB001934; BAA21621.1; JOINED.  
DR EMBL; AB001935; BAA21621.1; JOINED.  
DR EMBL; AB001936; BAA21621.1; JOINED.  
DR EMBL; AB001937; BAA21621.1; JOINED.  
DR EMBL; AB001938; BAA21621.1; JOINED.  
DR EMBL; AB001939; BAA21621.1; JOINED.  
DR EMBL; AB001940; BAA21621.1; JOINED.  
DR EMBL; AB001941; BAA21621.1; JOINED.  
DR EMBL; AB001942; BAA21621.1; JOINED.  
DR EMBL; AB001943; BAA21621.1; JOINED.  
DR EMBL; AB001944; BAA21621.1; JOINED.  
DR EMBL; AB001945; BAA21621.1; JOINED.  
DR EMBL; AB001946; BAA21621.1; JOINED.  
DR EMBL; AB001947; BAA21621.1; JOINED.  
DR EMBL; AB001948; BAA21621.1; JOINED.  
DR EMBL; AB001949; BAA21621.1; JOINED.  
DR EMBL; AB001950; BAA21621.1; JOINED.  
DR EMBL; AB001951; BAA21621.1; JOINED.  
DR EMBL; AB001952; BAA21621.1; JOINED.  
DR EMBL; AB001953; BAA21621.1; JOINED.  
DR EMBL; AB001954; BAA21621.1; JOINED.  
DR EMBL; AB001955; BAA21621.1; JOINED.  
DR EMBL; AB001956; BAA21621.1; JOINED.  
DR EMBL; AB001957; BAA21621.1; JOINED.  
DR EMBL; AB001958; BAA21621.1; JOINED.  
DR EMBL; AB001959; BAA21621.1; JOINED.  
DR EMBL; AB001960; BAA21621.1; JOINED.  
DR EMBL; AB001961; BAA21621.1; JOINED.  
DR EMBL; AB001962; BAA21621.1; JOINED.  
DR EMBL; AB001963; BAA21621.1; JOINED.  
DR EMBL; AB001964; BAA21621.1; JOINED.  
DR EMBL; AB001965; BAA21621.1; JOINED.  
DR EMBL; AB001966; BAA21621.1; JOINED.  
DR EMBL; AB001967; BAA21621.1; JOINED.  
DR EMBL; AB001968; BAA21621.1; JOINED.  
DR EMBL; AB001969; BAA21621.1; JOINED.  
DR EMBL; AB001970; BAA21621.1; JOINED.  
DR EMBL; AB001971; BAA21621.1; JOINED.  
DR EMBL; AB001972; BAA21621.1; JOINED.  
DR EMBL; AB001973; BAA21621.1; JOINED.  
DR EMBL; AB001974; BAA21621.1; JOINED.  
DR EMBL; AB001975; BAA21621.1; JOINED.  
DR EMBL; AB001976; BAA21621.1; JOINED.  
DR EMBL; AB001977; BAA21621.1; JOINED.  
DR EMBL; AB001978; BAA21621.1; JOINED.  
DR EMBL; AB001979; BAA21621.1; JOINED.  
DR EMBL; AB001980; BAA21621.1; JOINED.  
DR EMBL; AB001981; BAA21621.1; JOINED.  
DR EMBL; AB001982; BAA21621.1; JOINED.  
DR EMBL; AB001983; BAA21621.1; JOINED.  
DR EMBL; AB001984; BAA21621.1; JOINED.  
DR EMBL; AB001985; BAA21621.1; JOINED.  
DR EMBL; AB001986; BAA21621.1; JOINED.  
DR EMBL; AB001987; BAA21621.1; JOINED.  
DR EMBL; AB001988; BAA21621.1; JOINED.  
DR EMBL; AB001989; BAA21621.1; JOINED.  
DR EMBL; AB001990; BAA21621.1; JOINED.  
DR EMBL; AB001991; BAA21621.1; JOINED.  
DR EMBL; AB001992; BAA21621.1; JOINED.  
DR EMBL; AB001993; BAA21621.1; JOINED.  
DR EMBL; AB001994; BAA21621.1; JOINED.  
DR EMBL; AB001995; BAA21621.1; JOINED.  
DR EMBL; AB001996; BAA21621.1; JOINED.  
DR EMBL; AB001997; BAA21621.1; JOINED.  
DR EMBL; AB001998; BAA21621.1; JOINED.  
DR EMBL; AB001999; BAA21621.1; JOINED.  
DR EMBL; AB002000; BAA21621.1; JOINED.  
DR EMBL; AB002001; BAA21621.1; JOINED.  
DR EMBL; AB002002; BAA21621.1; JOINED.  
DR EMBL; AB002003; BAA21621.1; JOINED.  
DR EMBL; AB002004; BAA21621.1; JOINED.  
DR EMBL; AB002005; BAA21621.1; JOINED.  
DR EMBL; AB002006; BAA21621.1; JOINED.  
DR EMBL; AB002007; BAA21621.1; JOINED.  
DR EMBL; AB002008; BAA21621.1; JOINED.  
DR EMBL; AB002009; BAA21621.1; JOINED.  
DR EMBL; AB002010; BAA21621.1; JOINED.  
DR EMBL; AB002011; BAA21621.1; JOINED.  
DR EMBL; AB002012; BAA21621.1; JOINED.  
DR EMBL; AB002013; BAA21621.1; JOINED.  
DR EMBL; AB002014; BAA21621.1; JOINED.  
DR EMBL; AB002015; BAA21621.1; JOINED.  
DR EMBL; AB002016; BAA21621.1; JOINED.  
DR EMBL; AB002017; BAA21621.1; JOINED.  
DR EMBL; AB002018; BAA21621.1; JOINED.  
DR EMBL; AB002019; BAA21621.1; JOINED.  
DR EMBL; AB002020; BAA21621.1; JOINED.  
DR EMBL; AB002021; BAA21621.1; JOINED.  
DR EMBL; AB002022; BAA21621.1; JOINED.  
DR EMBL; AB002023; BAA21621.1; JOINED.  
DR EMBL; AB002024; BAA21621.1; JOINED.  
DR EMBL; AB002025; BAA21621.1; JOINED.  
DR EMBL; AB002026; BAA21621.1; JOINED.  
DR EMBL; AB002027; BAA21621.1; JOINED.  
DR EMBL; AB002028; BAA21621.1; JOINED.  
DR EMBL; AB002029; BAA21621.1; JOINED.  
DR EMBL; AB002030; BAA21621.1; JOINED.  
DR EMBL; AB002031; BAA21621.1; JOINED.  
DR EMBL; AB002032; BAA21621.1; JOINED.  
DR EMBL; AB002033; BAA21621.1; JOINED.  
DR EMBL; AB002034; BAA21621.1; JOINED.  
DR EMBL; AB002035; BAA21621.1; JOINED.  
DR EMBL; AB002036; BAA21621.1; JOINED.  
DR EMBL; AB002037; BAA21621.1; JOINED.  
DR EMBL; AB002038; BAA21621.1; JOINED.  
DR EMBL; AB002039; BAA21621.1; JOINED.  
DR EMBL; AB002040; BAA21621.1; JOINED.  
DR EMBL; AB002041; BAA21621.1; JOINED.  
DR EMBL; AB002042; BAA21621.1; JOINED.  
DR EMBL; AB002043; BAA21621.1; JOINED.  
DR EMBL; AB002044; BAA21621.1; JOINED.  
DR EMBL; AB002045; BAA21621.1; JOINED.  
DR EMBL; AB002046; BAA21621.1; JOINED.  
DR EMBL; AB002047; BAA21621.1; JOINED.  
DR EMBL; AB002048; BAA21621.1; JOINED.  
DR EMBL; AB002049; BAA21621.1; JOINED.  
DR EMBL; AB002050; BAA21621.1; JOINED.  
DR EMBL; AB002051; BAA21621.1; JOINED.  
DR EMBL; AB002052; BAA21621.1; JOINED.  
DR EMBL; AB002053; BAA21621.1; JOINED.  
DR EMBL; AB002054; BAA21621.1; JOINED.  
DR EMBL; AB002055; BAA21621.1; JOINED.  
DR EMBL; AB002056; BAA21621.1; JOINED.  
DR EMBL; AB002057; BAA21621.1; JOINED.  
DR EMBL; AB002058; BAA21621.1; JOINED.  
DR EMBL; AB002059; BAA21621.1; JOINED.  
DR EMBL; AB002060; BAA21621.1; JOINED.  
DR EMBL; AB002061; BAA21621.1; JOINED.  
DR EMBL; AB002062; BAA21621.1; JOINED.  
DR EMBL; AB002063; BAA21621.1; JOINED.  
DR EMBL; AB002064; BAA21621.1; JOINED.  
DR EMBL; AB002065; BAA21621.1; JOINED.  
DR EMBL; AB002066; BAA21621.1; JOINED.  
DR EMBL; AB002067; BAA21621.1; JOINED.  
DR EMBL; AB002068; BAA21621.1; JOINED.  
DR EMBL; AB002069; BAA21621.1; JOINED.  
DR EMBL; AB002070; BAA21621.1; JOINED.  
DR EMBL; AB002071; BAA21621.1; JOINED.  
DR EMBL; AB002072; BAA21621.1; JOINED.  
DR EMBL; AB002073; BAA21621.1; JOINED.  
DR EMBL; AB002074; BAA21621.1; JOINED.  
DR EMBL; AB002075; BAA21621.1; JOINED.  
DR EMBL; AB002076; BAA21621.1; JOINED.  
DR EMBL; AB002077; BAA21621.1; JOINED.  
DR EMBL; AB002078; BAA21621.1; JOINED.  
DR EMBL; AB002079; BAA21621.1; JOINED.  
DR EMBL; AB002080; BAA21621.1; JOINED.  
DR EMBL; AB002081; BAA21621.1; JOINED.  
DR EMBL; AB002082; BAA21621.1; JOINED.  
DR EMBL; AB002083; BAA21621.1; JOINED.  
DR EMBL; AB002084; BAA21621.1; JOINED.  
DR EMBL; AB002085; BAA21621.1; JOINED.  
DR EMBL; AB002086; BAA21621.1; JOINED.  
DR EMBL; AB002087; BAA21621.1; JOINED.  
DR EMBL; AB002088; BAA21621.1; JOINED.  
DR EMBL; AB002089; BAA21621.1; JOINED.  
DR EMBL; AB002090; BAA21621.1; JOINED.  
DR EMBL; AB002091; BAA21621.1; JOINED.  
DR EMBL; AB002092; BAA21621.1; JOINED.  
DR EMBL; AB002093; BAA21621.1; JOINED.  
DR EMBL; AB002094; BAA21621.1; JOINED.  
DR EMBL; AB002095; BAA21621.1; JOINED.  
DR EMBL; AB002096; BAA21621.1; JOINED.  
DR EMBL; AB002097; BAA21621.1; JOINED.  
DR EMBL; AB002098; BAA21621.1; JOINED.  
DR EMBL; AB002099; BAA21621.1; JOINED.  
DR EMBL; AB002100; BAA21621.1; JOINED.  
DR EMBL; AB002101; BAA21621.1; JOINED.  
DR EMBL; AB002102; BAA21621.1; JOINED.  
DR EMBL; AB002103; BAA21621.1; JOINED.  
DR EMBL; AB002104; BAA21621.1; JOINED.  
DR EMBL; AB002105; BAA21621.1; JOINED.  
DR EMBL; AB002106; BAA21621.1; JOINED.  
DR EMBL; AB002107; BAA21621.1; JOINED.  
DR EMBL; AB002108; BAA21621.1; JOINED.  
DR EMBL; AB002109; BAA21621.1; JOINED.  
DR EMBL; AB002110; BAA21621.1; JOINED.  
DR EMBL; AB002111; BAA21621.1; JOINED.  
DR EMBL; AB002112; BAA21621.1; JOINED.  
DR EMBL; AB002113; BAA21621.1; JOINED.  
DR EMBL; AB002114; BAA21621.1; JOINED.  
DR EMBL; AB002115; BAA21621.1; JOINED.  
DR EMBL; AB002116; BAA21621.1; JOINED.  
DR EMBL; AB002117; BAA21621.1; JOINED.  
DR EMBL; AB002118; BAA21621.1; JOINED.  
DR EMBL; AB002119; BAA21621.1; JOINED.  
DR EMBL; AB002120; BAA21621.1; JOINED.  
DR EMBL; AB002121; BAA21621.1; JOINED.  
DR EMBL; AB002122; BAA21621.1; JOINED.  
DR EMBL; AB002123; BAA21621.1; JOINED.  
DR EMBL; AB002124; BAA21621.1; JOINED.  
DR EMBL; AB002125; BAA21621.1; JOINED.  
DR EMBL; AB002126; BAA21621.1; JOINED.  
DR EMBL; AB002127; BAA21621.1; JOINED.  
DR EMBL; AB002128; BAA21621.1; JOINED.  
DR EMBL; AB002129; BAA21621.1; JOINED.  
DR EMBL; AB002130; BAA21621.1; JOINED.  
DR EMBL; AB002131; BAA21621.1; JOINED.  
DR EMBL; AB002132; BAA21621.1; JOINED.  
DR EMBL; AB002133; BAA21621.1; JOINED.  
DR EMBL; AB002134; BAA21621.1; JOINED.  
DR EMBL; AB002135; BAA21621.1; JOINED.  
DR EMBL; AB002136; BAA21621.1; JOINED.  
DR EMBL; AB002137; BAA21621.1; JOINED.  
DR EMBL; AB002138; BAA21621.1; JOINED.  
DR EMBL; AB002139; BAA21621.1; JOINED.  
DR EMBL; AB002140; BAA21621.1; JOINED.  
DR EMBL; AB002141; BAA21621.1; JOINED.  
DR EMBL; AB002142; BAA21621.1; JOINED.  
DR EMBL; AB002143; BAA21621.1; JOINED.  
DR EMBL; AB002144; BAA21621.1; JOINED.  
DR EMBL; AB002145; BAA21621.1; JOINED.  
DR EMBL; AB002146; BAA21621.1; JOINED.  
DR EMBL; AB002147; BAA21621.1; JOINED.  
DR EMBL; AB002148; BAA21621.1; JOINED.  
DR EMBL; AB002149; BAA21621.1; JOINED.  
DR EMBL; AB002150; BAA21621.1; JOINED.  
DR EMBL; AB002151; BAA21621.1; JOINED.  
DR EMBL; AB002152; BAA21621.1; JOINED.  
DR EMBL; AB002153; BAA21621.1; JOINED.  
DR EMBL; AB002154; BAA21621.1; JOINED.  
DR EMBL; AB002155; BAA21621.1; JOINED.  
DR EMBL; AB002156; BAA21621.1; JOINED.  
DR EMBL; AB002157; BAA21621.1; JOINED.  
DR EMBL; AB002158; BAA21621.1; JOINED.  
DR EMBL; AB002159; BAA21621.1; JOINED.  
DR EMBL; AB002160; BAA21621.1; JOINED.  
DR EMBL; AB002161; BAA21621.1; JOINED.  
DR EMBL; AB002162; BAA21621.1; JOINED.  
DR EMBL; AB002163; BAA21621.1; JOINED.  
DR EMBL; AB002164; BAA21621.1; JOINED.  
DR EMBL; AB002165; BAA21621.1; JOINED.  
DR EMBL; AB002166; BAA21621.1; JOINED.  
DR EMBL; AB002167; BAA21621.1; JOINED.  
DR EMBL; AB002168; BAA21621.1; JOINED.  
DR EMBL; AB002169; BAA21621.1; JOINED.  
DR EMBL; AB002170; BAA21621.1; JOINED.  
DR EMBL; AB002171; BAA21621.1; JOINED.  
DR EMBL; AB002172; BAA21621.1; JOINED.  
DR EMBL; AB002173; BAA21621.1; JOINED.  
DR EMBL; AB002174; BAA21621.1; JOINED.  
DR EMBL; AB002175; BAA21621.1; JOINED.  
DR EMBL; AB002176; BAA21621.1; JOINED.  
DR EMBL; AB002177; BAA21621.1; JOINED.  
DR EMBL; AB002178; BAA21621.1; JOINED.  
DR EMBL; AB002179; BAA21621.1; JOINED.  
DR EMBL; AB002180; BAA21621.1; JOINED.  
DR EMBL; AB002181; BAA21621.1; JOINED.  
DR EMBL; AB002182; BAA21621.1; JOINED.  
DR EMBL; AB002183; BAA21621.1; JOINED.  
DR EMBL; AB002184; BAA21621.1; JOINED.  
DR EMBL; AB002185; BAA21621.1; JOINED.  
DR EMBL; AB002186; BAA21621.1; JOINED.  
DR EMBL; AB002187; BAA21621.1; JOINED.  
DR EMBL; AB002188; BAA21621.1; JOINED.  
DR EMBL; AB002189; BAA21621.1; JOINED.  
DR EMBL; AB002190; BAA21621.1; JOINED.  
DR EMBL; AB002191; BAA21621.1; JOINED.  
DR EMBL; AB002192; BAA21621.1; JOINED.  
DR EMBL; AB002193; BAA21621.1; JOINED.  
DR EMBL; AB002194; BAA21621.1; JOINED.  
DR EMBL; AB002195; BAA21621.1; JOINED.  
DR EMBL; AB002196; BAA21621.1; JOINED.  
DR EMBL; AB002197; BAA21621.1; JOINED.  
DR EMBL; AB002198; BAA21621.1; JOINED.  
DR EMBL; AB002199; BAA21621.1; JOINED.  
DR EMBL; AB002200; BAA21621.1; JOINED.  
DR EMBL; AB002201; BAA21621.1; JOINED.  
DR EMBL; AB002202; BAA21621.1; JOINED.  
DR EMBL; AB002203; BAA21621.1; JOINED.  
DR EMBL; AB002204; BAA21621.1; JOINED.  
DR EMBL; AB002205; BAA21621.1; JOINED.  
DR EMBL; AB002206; BAA21621.1; JOINED.  
DR EMBL; AB002207; BAA21621.1; JOINED.  
DR EMBL; AB002208; BAA21621.1; JOINED.  
DR EMBL; AB002209; BAA21621.1; JOINED.  
DR EMBL; AB002210; BAA21621.1; JOINED.  
DR EMBL; AB002211; BAA21621.1; JOINED.  
DR EMBL; AB002212; BAA21621.1; JOINED.  
DR EMBL; AB002213; BAA21621.1; JOINED.  
DR EMBL; AB002214; BAA21621.1; JOINED.  
DR EMBL; AB002215; BAA21621.1; JOINED.  
DR EMBL; AB002216; BAA21621.1; JOINED.  
DR EMBL; AB002217; BAA21621.1; JOINED.  
DR EMBL; AB002218; BAA21621.1; JOINED.  
DR EMBL; AB002219; BAA21621.1; JOINED.  
DR EMBL; AB002220; BAA21621.1; JOINED.  
DR EMBL; AB002221; BAA21621.1; JOINED.  
DR EMBL; AB002222; BAA21621.1; JOINED.  
DR EMBL; AB002223; BAA21621.1; JOINED.  
DR EMBL; AB002224; BAA21621.1; JOINED.  
DR EMBL; AB002225; BAA21621.1; JOINED.  
DR EMBL; AB002226; BAA21621.1; JOINED.  
DR EMBL; AB002227; BAA21621.1; JOINED.  
DR EMBL; AB002228; BAA21621.1; JOINED.  
DR EMBL; AB002229; BAA21621.1; JOINED.  
DR EMBL; AB002230; BAA21621.1; JOINED.  
DR EMBL; AB002231; BAA21621.1; JOINED.  
DR EMBL; AB002232; BAA21621.1; JOINED.  
DR EMBL; AB002233; BAA21621.1; JOINED.  
DR EMBL; AB002234; BAA21621.1; JOINED.  
DR EMBL; AB002235; BAA21621.1; JOINED.  
DR EMBL; AB002236; BAA21621.1; JOINED.  
DR EMBL; AB002237; BAA21621.1; JOINED.  
DR EMBL; AB002238; BAA21621.1; JOINED.  
DR EMBL; AB002239; BAA21621.1; JOINED.  
DR EMBL; AB002240; BAA21621.1; JOINED.  
DR EMBL; AB002241; BAA21621.1; JOINED.  
DR EMBL; AB002242; BAA21621.1; JOINED.  
DR EMBL; AB002243; BAA21621.1; JOINED.  
DR EMBL; AB002244; BAA21621.1; JOINED.  
DR EMBL; AB002245; BAA21621.1; JOINED.  
DR EMBL; AB002246; BAA21621.1; JOINED.  
DR EMBL; AB002247; BAA21621.1; JOINED.  
DR EMBL; AB002248; BAA21621.1; JOINED.  
DR EMBL; AB002249; BAA21621.1; JOINED.  
DR EMBL; AB002250; BAA21621.1; JOINED.  
DR EMBL; AB002251; BAA21621.1; JOINED.  
DR EMBL; AB002252; BAA21621.1; JOINED.  
DR EMBL; AB002253; BAA21621.1; JOINED.  
DR EMBL; AB002254; BAA21621.1; JOINED.  
DR EMBL; AB002255; BAA21621.1; JOINED.  
DR EMBL; AB002256; BAA21621.1; JOINED.  
DR EMBL; AB002257; BAA21621.1; JOINED.  
DR EMBL; AB002258; BAA21621.1; JOINED.  
DR EMBL; AB002259; BAA21621.1; JOINED.  
DR EMBL; AB002260; BAA21621.1; JOINED.  
DR EMBL; AB002261; BAA21621.1; JOINED.  
DR EMBL; AB002262; BAA21621.1; JOINED.  
DR EMBL; AB002263; BAA21621.1; JOINED.  
DR EMBL; AB002264; BAA21621.1; JOINED.  
DR EMBL; AB002265; BAA21621.1; JOINED.  
DR EMBL; AB002266; BAA21621.1; JOINED.  
DR EMBL; AB002267; BAA21621.1; JOINED.  
DR EMBL; AB002268; BAA21621.1; JOINED.  
DR EMBL; AB002269; BAA21621.1; JOINED.  
DR EMBL; AB002270; BAA21621.1; JOINED.  
DR EMBL; AB002271; BAA21621.1; JOINED.  
DR E



```

FT CARBOHYD 1732 1732 N-LINKED (GLCNAC... ) (POTENTIAL).
FT VARSPPLIC 878 915 XXXXXXXXXXXXXXXXXXXXXXXXXXXX ->
FT ATERSWAEGECMLVKNNNLQORVLOQLCCKTCTFOG
FT (in isoform PC5A).
FT VARSPPLIC 916 1877 /Frid-VSP_005440.
FT Missing (in isoform PC5A).
SO SEQUENCE 1877 AA; 207888 MW; 890935DC60534444 CRC64;

Query Match 8.0%; Score 76.5; DB 1; Length 1877;
Best Local Similarity 25.7%; Pred. No. 1.1e+02;
Matches 29; Conservative 11; Mismatches 32; Indels 41; Gaps 6;

OY 21 LQKLVENYDLKA--TKNDPVW-----GNDFTCVGTAAQNLNDEKNE-AW-- 66
Db 110 VKRTRKRDYDLRSRAOSTYFNDPKWPSMYMHGCSDNTHPC-----QSDMNIEGAMKR 160
OY 67 -FMFMNNAADVQHTPEKATPPDKMYGYNKENATITYQTEDEQVLTDLVAFSDN 118
Db 161 GYTGKNIIVTILDDGIERTHPDLMOY-----DALASCDVN 196

```

Search completed: August 1, 2003, 12:15:48  
 Job time : 14.4838 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 12.7771 Seconds  
(without alignments)  
1294.578 Million cell updates/sec

Title: US-10-087-195-6  
Perfect score: 952  
Sequence: 1 DKPYMADEANGEHQDAWKH.....NEYAAGLPVRDVYTSDCLEPE 172

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_76:\*  
2: pir1:\*  
3: pir2:\*  
4: pir3:\*  
5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	9.3	919	2	S37786
2	88	9.2	648	2	T47896
3	88	9.2	1237	2	AC1583
4	86.5	9.1	368	2	S75234
5	85.5	9.0	745	2	B70017
6	84	8.8	408	2	AF0334
7	83.5	8.8	133	2	B81171
8	83.5	8.8	745	2	H85048
9	82.5	8.7	450	2	H82891
10	82.5	8.7	475	2	AB6811
11	82	8.6	1668	2	AB0272
12	81.5	8.6	133	2	G81933
13	81.5	8.6	512	2	T19806
14	81.5	8.6	919	2	S28179
15	81.5	8.6	1377	2	D90538
16	80.5	8.5	246	2	AE1347
17	80.5	8.5	865	1	C64737
18	80	8.4	387	2	AD2335
19	80	8.4	462	1	YSBYC
20	80	8.4	601	2	S57962
21	80	8.4	932	2	I52527
22	79.5	8.4	2335	2	T40186
23	79	8.3	580	2	AF2858
24	79	8.3	580	2	D97635
25	78	8.3	921	2	AB3968
26	78	8.2	874	2	T19540
27	78	8.2	938	2	E90632
28	78	8.2	938	2	D85483
29	77.5	8.1	283	2	AE2569

30	77.5	8.1	487	2	B39490	subtilisin-like pr
31	77.5	8.1	515	2	H72455	hypothetical prote
32	77.5	8.1	652	1	JC2101	subtilisin-like pr
33	77.5	8.1	962	2	JC5571	subtilisin-like pr
34	77.5	8.1	969	1	A39490	subtilisin-like pr
35	77.5	8.1	975	2	JC5570	subtilisin-like pr
36	77.5	8.1	1157	2	F97255	fusion of alpha-g1
37	77	8.1	1544	2	H86828	fructuronate reduc
38	77	8.1	754	2	T18238	lysophospholipase
39	76.5	8.0	915	2	B48225	probable proproctel
40	76	8.0	298	2	F90495	conserved hypothet
41	76	8.0	387	2	D70090	hypothetical prote
42	76	8.0	672	1	S73715	cyathodence acces
43	75.5	7.9	246	2	A11717	hypothetical prote
44	75.5	7.9	327	2	D89931	branched-chain alp
45	75.5	7.9	461	1	HMHU7A	alpha-L-fucosidase

## ALIGNMENTS

## RESULT 1

S37786 hypothetical protein YKL165c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YKL619

C:Species: Saccharomyces cerevisiae

C>Date: 03-May-1994 #sequence revision 03-May-1994 #text\_change 29-Oct-1999

C:Accession: S37786; S37995; S44563

R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.

Submitted to the EMBL Data Library, September 1993

A:Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4

A:Reference number: S37786

A:Accession: S37786

A:Molecule type: DNA

A:Residues: 1-919 <V>A>

A:Cross-references: EMBL:Z26877; NID:g407482; PIDN:CA081489.1; PID:g407483

A:Experimental source: strain S288C

R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.

Submitted to the Protein Sequence Database, March 1994

A:Reference number: S37976

A:Accession: S37995

A:Molecule type: DNA

A:Residues: 1-919 <V>A>

A:Cross-references: EMBL:Z26877; NID:g486288; PIDN:CA082007.1; PID:g486289; MIPS:YKL1

A:Experimental source: strain S288C

R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.

Submitted to the FAS1 and LAP4 loci

A:Reference number: S44563

A:Accession: S44563

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-919 <V>A>

A:Cross-references: EMBL:Z26877; NID:g407482; PIDN:CA081489.1; PID:g407483

A:Experimental source: strain S288C

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199

C:Genetics:

A:Gene: SGD:MCD4

A:Cross-references: SGD:S0001648; MIPS:YKL165c

A:Map position: 111

Query Match 9.3%; Score 88.5; DB 2; Length 919;  
Best Local Similarity 20.1%; Pred. No. 7.2;  
Matches 34; Conservative 30; Mismatches 64; Indels 41; Gaps 7;

QY 27 ENYDIKATYKNDPVGNDFTCVGTAONLN-----EDEKVEAMFMFMNNAD 74  
DB 117 EDVSAVTGKMGKPNFNSFQSTHTYSFCSPLDPMFKGASDPNKNVDTW----- 168  
QY 75 TYVQHTFEKATPDKMGYGNKNAITYQTED---GQVLTDLVAF---SDDNCRYVYALGP 127  
DB 169 -MYDHTFEDFTQOST---EIDDAVEYFRILDOLFHNSTLNSTLDYEIRODGNVFLHLGCG 223



Qy 128 DSGAGYELMATDITD-----VPASCLEKENEYAGLPVRYTSD 168  
Db 224 DFGHSHYRPSAEYDNDWKYIDQIPILIDKVKKFFADDTAFITAD 271

## RESULT 2

T47896

hypothetical protein T4C21.260 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 05-May-2000  
C:Accession: T47896  
R:Choi, N.; Robert, C.; Brotlier, P.; Wincker, P.; Catolico, L.; Attiguenave, F.; S.  
W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
Submitted to the Protein Sequence Database, March 2000  
A:Reference number: 224479  
A:Accession: T47896  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-648 <CHO>  
A:Cross-references: EMBL:AL162295  
A:Experimental source: cultivar Columbia; BAC clone T4C21  
C:Genetics:  
A:Map position: 3  
A:Note: T4C21.260  
C:Superfamily: Arabidopsis thaliana hypothetical protein T4C21.260

## Query Match

9.2%; Score 88; DB 2; Length 648;  
Best Local Similarity 25.3%; Pred. No. 5.1;

Matches 38; Conservative 19; Mismatches 69; Indels 24; Gaps 6;

Qy 2 KPWADPAANGEHODANKHLOKVEENYDLIKATYKNDPVGNDFTCVGTAONLNDEK 61  
Db 146 RAWKMDNEETAYLTAKRKID-----KLSRLDPATGNEFLCF---KNSQOFVS 193  
Qy 62 NVEAM--FMFMNNAADYVYOHTEFEKATPDKMYGYNKENAYIQFEDGQVLDVLAFSDDNC 119  
Db 194 HINOMODIYVGFADGKLGLGOKETIDRKQVWTGAKLYGIPESFIRCVAA-----C 248  
Qy 120 VYVYALGPDGSGAGYELMATDITD---DVPA 146  
Db 249 PVCNA--DSGSASRNKRREFYESLDVPA 276

## RESULT 3

AC1583

internalin protein (IPXG motif) homolog lin1204 [imported] - Listeria innocua (strain C  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AC1583  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Eschl, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitounnam, A.; Ma  
ok, C.; Schluster, T.; Simes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,  
A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1583  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1237 <GIA>  
A:Cross-references: GB:AL52022; PIDN:GAC96435.1; PID:g16413678; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lin1204

## Query Match

9.2%; Score 88; DB 2; Length 1237;  
Best Local Similarity 24.8%; Pred. No. 12;

Matches 41; Conservative 23; Mismatches 65; Indels 36; Gaps 6;

Qy 19 KHLQKVEENYDLIKATYKNDPVGNDFT-----CVGTAONLN 57  
Db 775 KYIVTLVNERLMDIKATPEVYIVDAVAGANITVYEDSGNKLAEINSTLTGNVGEYS 834

Qy 58 EDEKNVAMFNMNADYVYOHTEFEKATPDKMYGN-----KENATYQTEDEGOVL--TDV 111  
Db 835 SSAKEILGYLTLEPNSNAQGEFSLBEGTIVYIYSKKNVPKADITVQYTDGDELATET 894

Qy 112 LAESDNCYIYVYALGPDGSGAGYELMATDITDVPASCLEKENEY 156  
Db 895 LSGNVDENYVYTA---KTFGYEL-----IEPNSAEGFSEMA 930

## RESULT 4

S75234

hypothetical protein slr1920 - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S75234  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S75234  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-368 <RAN>  
A:Cross-references: EMBL:ID90903; GB:AB001339; NID:g1652127; PIDN:BA017148.1; PID:g165  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: Synechocystis hypothetical protein slr1920

## Query Match

9.1%; Score 86.5; DB 2; Length 368;  
Best Local Similarity 26.2%; Pred. No. 3.4;

Matches 43; Conservative 24; Mismatches 50; Indels 47; Gaps 10;

Qy 22 QKLEENYDL-----IKATYKNDPVGNDFTCVGTAONLNDEKNEAMFMNNA 73  
Db 112 QKLYRESFNLOPGLVAIKDIRKSIVSEKDY-----ADNIDK-----FLEIFN- 156  
Qy 74 DTVYQ-----HTFEKATPDKMYGYNKENAYIQFEDGQVLDVLAFSDDNCYIYALGPDG 129  
Db 157 PAIHQILLIRFIPPPVDDMHQDN---ILPPOFEKITLDVINFCIDLN---YSL----- 205  
Qy 130 SGAGYELM-----ATDVTVPASCLEKENEYAGLPVRYTSD 168  
Db 206 DNTFYDKWALVHALDETGVSLNVNKHQY-----PPREFETD 245

## RESULT 5

B70017

probable oxidoreductase (EC 1.-.-.-) yurc [similarity] - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Dec-2002  
C:Accession: B70017  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbre, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal  
lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Teipstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
A:Reference number: A65580; MUID:98044033; PMID:9384377  
A:Accession: B70017  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA



A:Residues: 1-745 <KUN>  
A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15238.1; PID:g2635745  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yurC  
A:Superfamily: carbon-monoxide dehydrogenase molybdoprotein  
C:Keywords: oxidoreductase

Query Match 9.0%; Score 85.5; DB 2; Length 745;  
Best Local Similarity 26.2%; Pred. No. 10;  
Matches 32; Conservative 22; Mismatches 41; Indels 27; Gaps 6;

OY 22 QKLEENYDLIKATYKNDPYWGNDFCTGTAQNLNEDKNEVAMFMNN-----A 73  
DB 106 EELAEALLEIQLVEYKELEWMDSPKALRPNAQRLHEDGNLHNAF-FSNDVBEGRAS 164  
OY 74 DTYVQHTFEKATPDKMGYKNEKNAITYQTEGQVLTDLVAFSDNCTVYIALGPDGSGAG 133  
DB 165 DTYVEETVE--LPRQMHY-----METEGVAVPE---DDGFTMYA---GTQHG 206  
OY 134 YE 135  
DB 207 YK 208

RESULT 6  
AF0334  
probable cytochrome c-type biogenesis protein YPO2743 [imported] - Yersinia pestis (stra

C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
A:Accession: AF0334

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AF0334

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-408 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC92981.1; PID:g15980720; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO2743

Query Match 8.8%; Score 84; DB 2; Length 408;  
Best Local Similarity 24.1%; Pred. No. 6.5;  
Matches 39; Conservative 21; Mismatches 60; Indels 42; Gaps 8;

OY 13 EHODAMKHLQKLEENYDLIKATYKNDPYWGNDFCTGTAQNLNEDKNEVAMFM 68  
DB 114 EGVAAHNOVAQME-----LRARVANEHADRLMEDVARLGLRTSLQDAGVNDMM 169  
OY 69 F-----MNAADTVQ---HTFEKATPDK---MYGKNEKNAITYQTEGQVLTDLVAFSD 116  
DB 170 LGRVGMALNATATQAFNAHYQLA-PDNNNEVKLGVEVLTFRSDPBNQATQMLR--- 225  
OY 117 DNCVYIALGPDGSGAGYELMATDYTDVPASCLEKFEHYAG 158  
DB 226 -----AMYGODHTMLRAVSLAFNAFEGG 249

RESULT 7

B81171

hypothetical protein MMB0676 [imported] - Neisseria meningitidis (strain MC58 serogroup

C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: B81171

R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
11, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Va

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: B81171

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-133 <TEP>

A:Cross-references: GB:AE002422; GB:AE002098; NID:g7225898; PIDN:AAF41094.1; PID:g722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: MMB0676

Query Match 8.8%; Score 83.5; DB 2; Length 133;  
Best Local Similarity 25.5%; Pred. No. 1.7;  
Matches 42; Conservative 17; Mismatches 53; Indels 53; Gaps 10;

OY 13 EHODAMKHLQKLEENYDLIKATYKNDPYWGNDFCTGTAQNLNEDKNEVAMFM 64  
DB 3 QHLEKWOHLNREEQKILAE-----VGLVQNDQEVHYEMKLKMAPDEASGE 49  
OY 65 AMFMNNADTVYQHTFEKATPDKMGYKNEKNAITYQTEGQVLTDLVAFS---DQNCYV 121  
DB 50 FWFPMATLSTL-----PPNRSLG-----LKNNGRLATVAVSILSVATEDN--- 90  
OY 122 IYALGPDGSGAGYELMATDYTDVPASCLEKFEHYAGLPVADVYT 166  
DB 91 -----PDIP-----QLMAQKITALNYSGP-NLNGYVAVSPCTRICT 125

RESULT 8

H85048

probable transposon protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
A:Accession: H85048

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: H85048

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-745 <STO>

A:Cross-references: GB:NC\_001268; NID:g7267141; PIDN:CAB80809.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4G03860

A:Map position: 4

Query Match 8.8%; Score 83.5; DB 2; Length 745;  
Best Local Similarity 25.6%; Pred. No. 16;  
Matches 42; Conservative 21; Mismatches 62; Indels 39; Gaps 10;

OY 7 DEANGENQ-DAKHLQKLEENYDLIKATYKNDPYWGNDFCTGTAQNLNED 59  
DB 240 DEDTETVEPESWMEKRLNKPDVEVVEEYIREFEKNND-FWGRYPCYETLADGLIED 298  
OY 60 EKNVEAMFMNNADTVYQHTFEKATPDKMGYKNEKNAITYQTEGQVLTDLVAFSDNC 119  
DB 299 -----VQHOFKCHLETLSY---PRATYKEERIEFLS-TLDLSIKKL 337  
OY 120 VYIYALGPDGSGAG-----YELMATDYTDVPA-SCLEKFE 154  
DB 338 EELFGF-PSGNGTKPKFERKELKDLMATSGNNVPLNSARSKSNQ 380

RESULT 9

H82891

hypothetical protein U0441 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: H82891

R:Glass, J.I.; Letkovitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.  
submitted to Genbank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A:Reference number: A82870  
A:Accession: H82891  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-450 <GAD>  
A:Cross-references: GB:AE002140; GB:AF222894; MID:96899420; PIDN:AAF30853.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: U0441  
A:Genetic code: SGC3

Query Match 8.7%; Score 82.5; DB 2; Length 450;  
Best Local Similarity 22.2%; Pred. No. 10;  
Matches 28; Conservative 25; Mismatches 44; Indels 29; Gaps 6;

QY 7 DEANGSHODAMKHLQKLVENYDLIKATYKNDPVWGNDPTCVG---TAQNINED--- 59  
DB 19 DENSESHIKIMKELDLONDKIINSHGMSNP---NDLNLGTFNFKELINDKDF 75  
QY 60 EKAVEAMFEMNAD---TVQHTPEKATPDKM-----YGNKENATYYQT 102  
DB 76 KRELKXIDTLNMDIYNNTINKYKFKNYIDFISYFDIINIFSAGNSYND--IKYRN 133  
QY 103 EDGOVL 108  
DB 134 KNGHIL 139

RESULT 10  
A86811  
beta-glucosidase (EC 3.2.1.21) [imported] - Lactococcus lactis subsp. lactis (strain IL1  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: A86811  
R:Holcfin, A.; Wancher, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrli  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86825; MID:21235186; PMID:11337471  
A:Accession: A86811  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-475 <STO>  
A:Cross-references: GB:AE005176; PID:912724484; PIDN:AAK05587.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: YPCA  
C:Superfamily: Agrobacterium beta-glucosidase  
C:Keywords: glycosidase; hydrolase

Query Match 8.7%; Score 82.5; DB 2; Length 475;  
Best Local Similarity 17.9%; Pred. No. 11;  
Matches 34; Conservative 29; Mismatches 54; Indels 73; Gaps 8;

QY 4 VMADEANGSHODAMKHLQKLVENYDLIKATYKNDPVWGNDPTCVG--- 48  
DB 275 LMQSEA-----EDLKRKENKYDLGVNHYHSRVOEPYSSDLSAQMRPKYYA 325  
QY 49 -----VCTAQNINEDKENTYAMFEMNADTVYQHTPEKATPD 87  
DB 326 SFNKRCVRNADGMEIHFOITYEIKRIQEDYDGNISWFISEN----- 368  
QY 88 KMYGKKNATYYQTEGQVLDVY-LAFSDNCRYVALGPOSSGA-GYEIMATDVTDP 145  
DB 369 ---GKGVENHEHFKDEGQIQQDYRIITFTTEHLEHLMKAIQESNCGFVHW-----TP 419  
QY 146 ASCLKEFNEX 155  
DB 420 IDQWWRNNGY 429

RESULT 11  
A60272  
Iga-specific metalloendopeptidase (EC 3.4.24.13) - Streptococcus sanguis

N:Alternate names: Iga protease; immunoglobulin A1 protease  
C:Species: Streptococcus sanguis  
C>Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 21-Jan-2000  
C:Accession: A60272  
R:Gilbert, J.V.; Plant, A.G.; Wright, A.  
Infect. Immun. 59, 7-17, 1991  
A:Title: Analysis of the immunoglobulin A protease gene of Streptococcus sanguis.  
A:Reference number: A60272; MID:91100011; PMID:1987065  
A:Accession: A60272  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-1668 <GII>  
C:Comment: This protein is shown from the start of translation of this gene as determ  
the start codon is shown in entry B60272.  
C:Genetics:  
A:Gene: iga  
A:Start codon: GTG

Query Match 8.6%; Score 82; DB 2; Length 1668;  
Best Local Similarity 27.1%; Pred. No. 61;  
Matches 51; Conservative 12; Mismatches 47; Indels 78; Gaps 12;

QY 7 DEANGSHODAMKHLQKLVENYDLIKATYKNDPVWGNDPTCVG--- 1668  
DB 1160 DRYRNNHR-AGALNKFEVDAQETAKRORHRYLDROGREKLYRNILVYDAK 1218  
QY 38 NDPVWGNDPTCVGTAQNINEDKENTYAMFEMNAD-----DYYQHTPEKATPD 87  
DB 1219 ---FQDDTV-----DKATYEAQDSSNPMPKFFGVGKVVHNGKATATGD 1264  
QY 88 KMY-GY---NKENATY---QTEGQVLDVY-LAFSDNCRYVALGPOSSGA-GYEIM 137  
DB 1265 SYVYMGYRMDKRGATYTHETHD-----SDNEIY---LGGVGRSGGLCPERF 1310  
QY 138 ATTYTDP 145  
DB 1311 AKGLQAP 1318

RESULT 12  
G81933  
hypothetical protein MMA0876 [imported] - Neisseria meningitidis (strain Z2491 serogr  
C:Species: Neisseria meningitidis  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: G81933  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
A:Reference number: A81775; MID:2022556; PMID:10761919  
A:Accession: G81933  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-133 <PAR>  
A:Cross-references: GB:AL162754; GB:AL157959; MID:97379424; PIDN:CAB84156.1; PID:9737  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: MMA0876

Query Match 8.6%; Score 81.5; DB 2; Length 133;  
Best Local Similarity 24.8%; Pred. No. 2.6;  
Matches 41; Conservative 19; Mismatches 52; Indels 53; Gaps 10;

QY 13 EHODAMKHLQKLVENYDLIKATYKNDPVWG---NDFTCVGTAQNIN-EDKENV 64  
DB 3 QHIEKQHSREOKILAE-----VGLGVNDDEVHYEMAKLNAPEPASE 49  
QY 65 AMFEMNADTVYQHTPEKATPDKMYGNKENATYYQTEGQVLDVY-LAFS---DNCYV 121  
DB 50 FWRBAETLSTL-----PPNR-----SLDRMNGRLAFAVILSVIWDN--- 90



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 49.4705 Seconds  
(without alignments)  
897.203 Million cell updates/sec

Title: US-10-087-195-6  
Perfect score: 952  
Sequence: 1 DKPWADAEANGEHQDAKMH.....NEYAAGLPVRDVTSDCLPE 172

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	239	25.1	209	5	Q8WSK7
2	93.5	9.8	277	5	Q9W3B3
3	92.5	9.7	3198	5	Q9U8G8
4	92	9.7	2017	5	Q8ICR6
5	91	9.6	386	16	Q8XP49
6	90	9.5	1874	2	Q33765
7	88	9.2	648	10	Q9LXK9
8	88	9.2	677	2	Q8L2E8
9	88	9.2	1237	16	Q9ZC88
10	88	9.2	1854	2	Q33764
11	88	9.2	1854	2	Q33759
12	86.5	9.1	368	16	P73122
13	86	9.0	1898	5	Q8IL35
14	86	9.0	1928	5	Q9U0H2
15	85	8.9	1874	2	Q33763
16	84	8.8	408	16	Q8ZD50

17	83.5	8.8	133	16	Q9K0D5
18	83.5	8.7	745	10	Q9M105
19	83	8.7	757	10	Q9SRI4
20	82.5	8.7	441	16	Q8RAP3
21	82.5	8.7	450	16	Q9PQ49
22	82.5	8.7	475	16	Q9CFI7
23	82.5	8.7	688	3	Q9UV05
24	82	8.6	1853	2	Q33761
25	82	8.6	1854	2	Q33760
26	82	8.6	1878	2	Q33762
27	82	8.6	1914	2	Q33762
28	82	8.6	3896	5	Q8I581
29	81.5	8.6	133	16	Q9JVE2
30	81.5	8.6	388	10	Q8H750
31	81.5	8.6	391	5	Q9N6C3
32	81.5	8.6	512	5	Q62055
33	81.5	8.6	1377	16	Q98Q28
34	81.5	8.6	1798	16	Q9K211
35	81	8.5	754	3	Q9UVX1
36	81	8.5	1121	17	Q8TM44
37	81	8.5	1964	2	Q59947
38	80.5	8.5	246	16	Q8Y588
39	80.5	8.5	722	5	Q9U0H5
40	80.5	8.5	1462	12	Q8JVL1
41	80.5	8.5	1588	16	Q8KXK8
42	80	8.4	387	16	Q8YFP9
43	80	8.4	511	16	Q931Y5
44	80	8.4	601	2	Q45821
45	80	8.4	932	11	Q62030

#### ALIGNMENTS

##### RESULT 1

ID	Q8WSK7	PRELIMINARY;	PRT;	209 AA.
AC	Q8WSK7;			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Serotonin and histamine binding protein.			
OS	Dermacentor reticulatus.			
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
CC	Parasitiformes; Ixodidae; Ixodidae; Dermacentor.			
OX	NCBI_TaxID=57047;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Sangamadech S., Paesen G.C., Nuttall P.A.;			
RT	"A high affinity serotonin- and histamine-binding lipocalin secreted by blood-feeding ticks."			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF217101; AAL56644.1;			
DR	InterPro; IPR002970; His_binding.			
DR	Pfam; PF02098; His_binding; 1.			
DR	SEQUENCE 209 AA; 23889 MW; E5E88BE710859591 CRC64;			

QY	1	DKPWADAEANGEHQDAKMHQKVEENYDLKATYKNDP-VNGNDTCVGTAAQNND 59
DB	29	ENPLMAHEELIGKQDAKSIIDQSVYVIAKTYENDGWSQKCLQVQIEERKEE 88
QY	60	EKNVEAFMFNNADTV-YOTFEKATPDMMYGV-NKENATYQTEGGVLTFLAFSD- 116
DB	89	DYVTSVFTFRNASSPKIYNVETFAVAYOYKKNRNRIEYVGGLMTITDILFTDG 148
QY	117	DNCVYIYALPDGSGAGYELMA--TDYTDVPASCLEKFNYYAA-GLPVRDVTSDCL 170
DB	149	ELCDVFVPMND--QCELMVKKSHYKHVPDYCTFEVNFCAKRDRTYDIFNECV 202

OY 141 YTDVP 145  
Db 1544 LLOAP 1548

RESULT 7  
O9L2X9 PRELIMINARY; PRT; 648 AA.  
ID O9L2X9  
AC O9L2X9  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN T4C21\_260 OR A3G60850.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Choiane N., Robert C., Brotier P., Mincker P., Catolico L.,  
RA Attienave F., Saurin W., Weisenbach J., Mewes H.W., Rudd S.,  
RA Lemcke K., Mayer K.F.X., Quelet F., Salanoubat M.,  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
RA Deng J.M., Hayashizaki Y., Hsuan V.M., Lee J.M., Ishida J., Kamiya A.,  
RA Kawai J., Kim C.J., Narusaka M., Onodera C.S., Quach H.L., Sakurai T.,  
RA Satou M., Seki M., Shinozaki K., Toriumi M., Wong C., Wu H.C.,  
RA Yamaoka K., Yu G., Yuan S., Shinozaki K., Ecker J., Theologis A.,  
RA Davis R.M.,  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL162295; CAB82689.1;  
DR EMBL; BF002453; AAC00813.1;  
KM Hypothetical protein.  
SO SEQUENCE 648 AA; 72914 MW; 83A3F7551172FC58 CRC64;

Query Match 9.2%; Score 88; DB 10; Length 648;  
Best Local Similarity 25.3%; Pred. No. 23;  
Matches 38; Conservative 19; Mismatches 69; Indels 24; Gaps 6;

OY 2 KPYMADEANNGEHQDAKHLQKLVEDIKATYKNDPYWGDFTCVGTAAQNLNEDK 61  
Db 146 RAVMKRNETATVLTAMKRIOD-----KLTSLRDPATGNEFLCF--KNNSOQFVS 193  
OY 62 NVEAM--FMFMNADTVYOHTEFKATPKDKMYGKNEKATYYQEDCGVLTDLAFSDNC 119  
Db 194 HINQKODIYWGFIADGDKHLGKERTIDRKQVWYTGAKLYGIPESIRCVAA-----C 248  
OY 120 YVITVIALGPDGSGAGYELMATDYT--DVPA 146  
Db 249 PVCNA--DSGSASRNKRREPYTESLDVPA 276

RESULT 8  
O8L2E8 PRELIMINARY; PRT; 677 AA.  
ID O8L2E8  
AC O8L2E8  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Zinc metalloprotease Pap6.  
OS Vibrio Harveyi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrrio.  
OX NCBI\_TaxID=669;

RN [1]  
RP SEQUENCE FROM N.A.  
RA Teo J., Poh C.L., Zhang L.H.,  
RT "Vibrio harveyi zinc metalloprotease".  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF508306; AA34261.1;  
DR InterPro: IPR001570; Peptidase\_M4.  
DR InterPro: IPR005075; Pep\_M4\_propep.  
DR InterPro: IPR006025; Zn\_MTPeptidase.  
DR Pfam: PF01447; Peptidase\_M4; 1.  
DR Pfam: PF02968; Peptidase\_M4\_C; 1.  
DR Pfam: PF03413; Pep\_M4\_propep; 1.  
DR PRINTS: PR00730; THERMOLYSIN.  
DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
KM Protease; Metalloprotease.  
SO SEQUENCE 677 AA; 75120 MW; 5E904C0A127CA186 CRC64;

Query Match 9.2%; Score 88; DB 2; Length 677;  
Best Local Similarity 21.7%; Pred. No. 25;  
Matches 47; Conservative 26; Mismatches 60; Indels 84; Gaps 10;

OY 2 KPYMADEANNGEHQDAKHLQKLVEDIKATYKNDPYWGDFTCVGTAAQNLNEDK 61  
Db 169 RPTLVDAKSGELMDMEGLAFLEAGPGGNKSGRYTGPNTKFGGFVDYCCQNSDN 228  
OY 32 IKATYKNDPYWGD--FTCVGTAAQNLNEDKNEVAMFMFNADTVYOHTE----- 81  
Db 229 VETINNNNOOMGQVHRENC-----NVN-NYREINGAYAPNMDAHYFGQVDFMYREWL 281  
OY 82 -----EKATPKMYGKNEKATYYQEDCGVLT-----DYLA-----PS 115  
Db 282 GARPIQOKLRLMYRYSNGNAFW----DGRITTFGDSGSSMYPLATRDVIAHEVSHGFT 337  
OY 116 DDCVYVIALGPDGSGAGYELMATDYTDPASCLEKF 152  
Db 338 EQNSGLELR---GMSGWYE---SFSVAAALSLQ 367

RESULT 9  
O92CG8 PRELIMINARY; PRT; 1237 AA.  
ID O92CG8  
AC O92CG8  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Hypothetical protein lin1204.  
GN LIN1204.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CLIP 11262 / Serovar 6a;  
RX PubMed=11679669;  
RA Glaeser P., Frangeul L., Buchrieser C., Rusnock C., Amend A.,  
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Deboux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Ertlan K.-D., Esli H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,  
RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordstok G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.,  
RT "Comparative genomics of Listeria species.",  
RL Science 294:845-852(2001).  
DR EMBL; AL596168; CAC96435.1;  
DR Listlist: LIN01204;  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR006192; LPXTG.  
DR InterPro: IPR001611; LRR.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.



ID P73122 PRELIMINARY: PRT: 368 AA.  
AC P73122:  
DT 01-FEB-1997 (Tremblrel. 02, Created)  
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Hypothetical protein slr1920.  
GN SLR1920.  
OS *Synechocystis* sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.  
NCBI\_TaxID=1148;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Suglura M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT -Sequence analysis of the genome of the unicellular cyanobacterium  
RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL: D90903; BAA17148.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 368 AA; 43031 MW; 6F4612A3BDD17F3 CRC64;  
  
Query Match 9.1%; Score 86.5; DB 16; Length 368;  
Best Local Similarity 26.2%; Pred. No. 15;  
Matches 43; Conservative 24; Mismatches 50; Indels 47; Gaps 10;  
  
OY 22 OKLVEENYDL-----IKATYKNDPVGNDFCTGTAONLNDEKNEAMFMFMNNA 73  
DB 112 OKLYRESFNLQGHLLVAIKIRKKSIVGFKDY-----ADNLDK-----FLIEIFN- 156  
OY 74 DTVYQ---HTFEKATPDKMYGYNKRNATTTQTEGQVLTDLVAFSDNDCYIYALGPPG 129  
DB 157 PAIHOLLIRYFIPVDDMHQFDN--TELPPQEFKILTDVINEFDLN---YSL----- 205  
OY 130 SGAGYELW-----ATDYTPVAPASCLEKFNENYAGLFPVDVYTS 168  
DB 206 DNLITVQKQALVHALDETGVSLNLVKNHGEY-----PFRMEPTFD 245  
  
RESULT 13  
O81L35 PRELIMINARY: PRT: 1898 AA.  
AC O81L35:  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Hypothetical protein, conserved.  
GN PF14\_0414.  
OS *Plasmodium falciparum* (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; *Plasmodium*.  
NCBI\_TaxID=36329;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3D7;  
RX MEDLINE=22255705; PubMed=12368864;  
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,  
RA Chen M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,  
RA Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,  
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
RA Fraser C.M., Barrell B.;  
RT -Genome sequence of the human malaria parasite *Plasmodium*  
RT *falciparum*.";  
RL Nature 419:498-511(2002).  
DR EMBL: AE014823; AAN37027.1; -  
KW Hypothetical protein.

SQ SEQUENCE 1898 AA; 223829 MW; 7B5CB02A7091822F CRC64;  
  
Query Match 9.0%; Score 86; DB 5; Length 1898;  
Best Local Similarity 27.4%; Pred. No. 14e+02;  
Matches 31; Conservative 21; Mismatches 47; Indels 14; Gaps 6;  
  
OY 17 ANKHLOKVEENYDLIKATYKNDP---VGNDFCTGTAONLNDEKNEAMFMFMNNA 73  
DB 571 ANKNSVLEE--YCKSKYEKKKDKLLIDNYDNSLFTHYSNNHDSKEQSNFFPHNNA 627  
OY 74 DTVYQHTFEKATPDKM---YGYNK--ENATTTQTEGQVLTDLVAFSDNDCYI 121  
DB 628 NSSY-VTFDITDISENFISKYFPRISYKKRSYDFEGEKFDNLIL--GGMCYI 677  
  
RESULT 14  
O90UH2 PRELIMINARY: PRT: 1928 AA.  
AC O90UH2:  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
DE Hypothetical 229.7 kDa protein.  
GN MALP2\_50.  
OS *Plasmodium falciparum* (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; *Plasmodium*.  
NCBI\_TaxID=36329;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3D7;  
RA Devlin K., Bowman S., Churcher C., Harris B., Harris D., Lawson D.,  
RA Quail M., Barrell B.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL035475; CAB62891.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 1928 AA; 229692 MW; 38F6CECA9B3253B CRC64;  
  
Query Match 9.0%; Score 86; DB 5; Length 1928;  
Best Local Similarity 21.4%; Pred. No. 1.4e+02;  
Matches 33; Conservative 24; Mismatches 53; Indels 44; Gaps 6;  
  
OY 13 EHODANKHLQKVEENYDLIKATY-----KNDPVGNDF-TVGTGAONLNLD 59  
DB 702 DHVQMDKHSIAIINDIDINDVGTSCDNOVFPRIKNDVYDKFENCDELFEYENIQV 761  
OY 60 EKNV--EAMFMF--MNAADTVYQHTFEKATPDKMYGYNKRNAL----- 98  
DB 762 DONTYRKDFEFNDINNDHEITCENFSRYSLEKNEKEKIKNDKLIELSDIYSLKE 821  
OY 99 TQTEGQVLTDLVAFSDNDCYI 122  
DB 822 SYEGSEGVYIHNKIKSSDDTFSLGMNENNCYII 855  
  
RESULT 15  
O33763 PRELIMINARY: PRT: 1874 AA.  
AC O33763:  
ID O33763:  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE 1991 protease (EC 3.4.24.13).  
GN IGA.  
OS *Streptococcus sanguis*.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC *Streptococcus*.  
NCBI\_TaxID=1305;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SK4;  
RX MEDLINE=98084472; PubMed=9423856;  
RA Poulsen K., Reinholdt J., Jespersgaard C., Boye K., Brown T.A.,  
RA Hauge M., Kilian M.;







RT "A comprehensive genetic study of streptococcal immunoglobulin A1  
RT proteases: evidence for recombination within and between species.";  
RL Infect. Immun. 66:181-190(1998).  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY  
CC AN AMIDE BOND (BY SIMILARITY).  
DR EMBL: Y13459; CAA73856.1; -.  
DR MEROPS: M26.001; -.  
DR InterPro: IPR005877; Gpos\_Ysirk.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR006192; LPXTG.  
DR InterPro: IPR006025; Zn\_MTpeptidse.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR Pfam: PF04650; Ysirk\_signal; 1.  
DR TIGRFAMs: TIGR01167; LPXTG\_anchor; 1.  
DR TIGRFAMs: TIGR01168; Ysirk\_signal; 1.  
DR PROSITE: PS50847; Gram\_pos\_anchoring; 1.  
DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
KW Cell wall; Hydrolase; Peptidoglycan-anchor; Protease.  
SQ SEQUENCE 1874 AA; 207888 MW; FA30684926B84532 CRC64;

Query Match 8.98; Score 85; DB 2; Length 1874;

Best Local Similarity 27.68; Pred. No. 1,7e+02;  
Matches 51; Conservative 13; Mismatches 49; Indels 72; Gaps 13;

```
OY 7 DEANGSEHODAMKHLQKLVEN-----YD-----LIKATYK 37
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1390 DRYRNNNEHK-AGAELEKFEVDNAQETAKRQRDHDYRYRLDEQGREKLYRNILLYDAYK 1448
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 38 NDPVWGNDFTCVGTAQNLNDEKKNVEAMF-----MNNADTVYQHTFEKATPDPMY 90
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1449 ----FCDDAT-VGKATAEAQFDSSNPAMKFEFGVGNKVVHNRHGA-----ATGDGVY 1497
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 91 --GY---NKENATY---QTEDEQVLTDLAFSPDNCYVYIALGPDG--SGAGYELMATD 140
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1498 YMGYRMLDKDGAITYTHEMTHD-----SDNEIY---LGGYGRSGLGPEFFPAKG 1543
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 141 YTDVP 145
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1544 LLOAP 1548
```

Search completed: August 1, 2003, 12:19:57  
Job time : 51.4705 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:58 ; Search time 18.0267 Seconds

(without alignments)  
1199.017 Million cell updates/sec

Title: US-10-087-195-8

Perfect score: 989  
Sequence: 1 NPTMANEAKLSGYODAMKSL.....EKTIVRNYTDSCKPAPAQN 182

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Database :

Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpae/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpae/PCT\_NEM\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpae/US06\_NEM\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpae/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpae/US07\_NEM\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpae/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpae/US08\_NEM\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpae/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpae/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpae/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpae/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpae/US09\_NEM\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpae/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpae/US10C\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpae/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpae/US10\_NEM\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpae/US60\_PUBCOMB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpae/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	989	100.0	182	14	US-10-085-572-8		Sequence 8, Appl1
2	989	100.0	182	14	US-10-087-195-8		Sequence 8, Appl1
3	331.5	33.5	171	14	US-10-085-572-7		Sequence 7, Appl1
4	331.5	33.5	171	14	US-10-087-195-7		Sequence 7, Appl1
5	267.5	27.0	172	14	US-10-085-572-6		Sequence 6, Appl1
6	267.5	27.0	172	14	US-10-087-195-6		Sequence 6, Appl1
7	83.5	8.4	778	9	US-09-815-242-10728		Sequence 10728, A
8	79	8.0	221	9	US-09-728-914-4		Sequence 4, Appl1
9	76	7.7	2201	10	US-09-995-542-9		Sequence 9, Appl1
10	74.5	7.5	142	10	US-09-893-737-172		Sequence 172, App
11	74.5	7.5	198	9	US-09-864-761-48672		Sequence 48672, A
12	74.5	7.5	199	9	US-09-801-231-2		Sequence 2, Appl1
13	74.5	7.5	199	10	US-09-731-872-269		Sequence 269, App
14	74.5	7.5	199	14	US-10-052-586-588		Sequence 588, App
15	74.5	7.5	199	15	US-10-174-590-588		Sequence 588, App

16	74.5	7.5	199	15	US-10-176-758-588		Sequence 588, App
17	74.5	7.5	199	15	US-10-175-737-588		Sequence 588, App
18	74.5	7.5	199	15	US-10-173-706-588		Sequence 588, App
19	74.5	7.5	199	15	US-10-175-738-588		Sequence 588, App
20	74.5	7.5	199	15	US-10-175-752-588		Sequence 588, App
21	74.5	7.5	199	15	US-10-176-482-588		Sequence 588, App
22	74.5	7.5	199	15	US-10-176-757-588		Sequence 588, App
23	74.5	7.5	199	15	US-10-176-913-588		Sequence 588, App
24	74.5	7.5	199	15	US-10-180-532-588		Sequence 588, App
25	74.5	7.5	199	15	US-10-180-557-588		Sequence 588, App
26	74.5	7.5	199	15	US-10-173-700-588		Sequence 588, App
27	74.5	7.5	199	15	US-10-174-572-588		Sequence 588, App
28	74.5	7.5	199	15	US-10-174-579-588		Sequence 588, App
29	74.5	7.5	199	15	US-10-174-582-588		Sequence 588, App
30	74.5	7.5	199	15	US-10-174-588-588		Sequence 588, App
31	74.5	7.5	199	15	US-10-175-739-588		Sequence 588, App
32	74.5	7.5	199	15	US-10-175-740-588		Sequence 588, App
33	74.5	7.5	199	15	US-10-175-743-588		Sequence 588, App
34	74.5	7.5	199	15	US-10-176-488-588		Sequence 588, App
35	74.5	7.5	199	15	US-10-176-492-588		Sequence 588, App
36	74.5	7.5	199	15	US-10-176-747-588		Sequence 588, App
37	74.5	7.5	199	15	US-10-176-750-588		Sequence 588, App
38	74.5	7.5	199	15	US-10-176-985-588		Sequence 588, App
39	74.5	7.5	199	15	US-10-176-987-588		Sequence 588, App
40	74.5	7.5	199	15	US-10-176-992-588		Sequence 588, App
41	74.5	7.5	199	15	US-10-176-993-588		Sequence 588, App
42	74.5	7.5	199	15	US-10-184-658-588		Sequence 588, App
43	74.5	7.5	199	15	US-10-176-991-588		Sequence 588, App
44	74.5	7.5	199	15	US-10-173-695-588		Sequence 588, App
45	74.5	7.5	199	15	US-10-173-697-588		Sequence 588, App

## ALIGNMENTS

RESULT 1									
US-10-085-572-8									
Sequence 8, Application US/10085572									
Publication No. US20020151499A1									
GENERAL INFORMATION:									
APPLICANT: Nuttall, Patricia, Ann									
TITLE OF INVENTION: Treatment of Conjunctionitis									
FILE REFERENCE: 2488-1-003									
CURRENT APPLICATION NUMBER: US/10/085,572									
CURRENT FILING DATE: 2002-02-27									
PRIOR APPLICATION NUMBER: PCT/GB00/03282									
PRIOR FILING DATE: 2000-08-24									
PRIOR APPLICATION NUMBER: 9920674.0									
PRIOR FILING DATE: 1999-09-01									
NUMBER OF SEQ ID NOS: 8									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 8									
LENGTH: 182									
TYPE: PRT									
ORGANISM: MS-HBP 1									
US-10-085-572-8									
Query Match									
Best Local Similarity 100.0%; Score 989; DB 14; Length 182;									
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	NPTMANEAKLSGYODAMKSL	100.0%	Score 989;	DB 14;	Length 182;			
DB	1	NPTMANEAKLSGYODAMKSL	100.0%	Pred. No. 3.7e-99;					
QY	61	NATILKKKKHLTLKESHEITIVKAYDYTTENGKIKYETGRTOTFFEDVFVFSYKNC	100.0%						
DB	61	NATILKKKKHLTLKESHEITIVKAYDYTTENGKIKYETGRTOTFFEDVFVFSYKNC	100.0%						
QY	121	VIFVPEKSGDEGDEYELWVSEDKIDIPDCKFTMAVYPAQOQEKTVRNVYDSSCKPAPA	100.0%						
DB	121	VIFVPEKSGDEGDEYELWVSEDKIDIPDCKFTMAVYPAQOQEKTVRNVYDSSCKPAPA	100.0%						

QY 181 QN 182  
Db 181 QN 182

## RESULT 2

US-10-087-195-8  
; Sequence 8, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: MS-HBP 1  
US-10-087-195-8

Query Match 100.0%; Score 989; DB 14; Length 182;  
Best Local Similarity 100.0%; Pred. No. 3.7e-99;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPTWANEAKLSGYODAKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIGKKL 60  
Db 1 NPTWANEAKLSGYODAKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIGKKL 60  
QY 61 NATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFFVSDYKNC 120  
Db 61 NATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFFVSDYKNC 120  
QY 121 VIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVRNYTTDSSCKRPA 180  
Db 121 VIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVRNYTTDSSCKRPA 180  
QY 181 QN 182  
Db 181 QN 182

## RESULT 3

US-10-085-572-7  
; Sequence 7, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; PRIOR FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: FS-HBP 2  
US-10-085-572-7

Query Match 33.5%; Score 331.5; DB 14; Length 171;  
Best Local Similarity 42.0%; Pred. No. 5.3e-28;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;

QY 2 PTWANEAKLSGYODAKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIG--KKK 59  
Db 3 PDMADEANGAHQADAKSLKADVENYVWKATKNPVMQNDPTCGVAMANDVNEDEKS 62  
QY 60 LNATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFFVSDYKNC 119  
Db 63 IQAEFLFMNNAADTNMOPATEKVTAVKMYGYNRENAFRETED--GQVFTDVIAYSD--DNC 119  
QY 120 DVIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVRNYTTDSSC 175  
Db 120 DVIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVRNYTTDSSC 175

## RESULT 4

US-10-087-195-7  
; Sequence 7, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: FS-HBP 2  
US-10-087-195-7

Query Match 33.5%; Score 331.5; DB 14; Length 171;  
Best Local Similarity 42.0%; Pred. No. 5.3e-28;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;

QY 2 PTWANEAKLSGYODAKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIG--KKK 59  
Db 3 PDMADEANGAHQADAKSLKADVENYVWKATKNPVMQNDPTCGVAMANDVNEDEKS 62  
QY 60 LNATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFFVSDYKNC 119  
Db 63 IQAEFLFMNNAADTNMOPATEKVTAVKMYGYNRENAFRETED--GQVFTDVIAYSD--DNC 119  
QY 120 DVIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVRNYTTDSSC 175  
Db 120 DVIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVRNYTTDSSC 175

## RESULT 5

US-10-085-572-6  
; Sequence 6, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; PRIOR FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8



Query Match	8.0%	Score 79	DB 9	Length 221
Best Local Similarity	18.0%	Pred No. 1.77		
Matches	34	Conservative	26	Mismatches 61, Indels 68, Gaps 6
OY	29	YLAQATQTTDVGWGEFTCVSYTAERKIGKKKLATILYKKNHLTLDKESHETITV-WKAY	87	
		: : : : :   : : : : :   : : : : :   : : : : :		
Db	32	YGSCTTTTRPTGRCARMIVTTAPPEEDSK-----YKEQNATRVKEMNATQWKKRY	84	
OY	88	DYTEENGKIKYECOGRT-----QTEFDVY-----	112	
		: : : : :   : : : : :   : : : : :   : : : : :		
Db	85	DYTFDSGNPVCENFRVMEKRTPTVYSFOYRKKSNSMSTETDETLLADIGEHGPPNVNN	144	
OY	113	-----FSDYKNCQDVIYFVERKESDGDGYELAWS-----EDKIDKIP	148	
		: : : : :   : : : : :   : : : : :   : : : : :		
Db	145	FOPTPIGIATDNLVLYSNVNCVTRIP-FTNOSGRHODLMANNTLSQETFDCLNKEF	203	
OY	149	DCCKETMAY	157	
		: : : : :   : : : : :   : : : : :   : : : : :		
Db	204	EYCNTQIT	212	

```

RESULT 9
US-09-995-542-9
; Sequence 9, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ullas, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (115)
; OTHER INFORMATION: amino acid at this position is unknown
; US-09-995-542-9

```

```

Query Match      7.7%; Score 76; DB 10; length 2201;
Best Local Similarity 24.1%; Pred. No. 85;
Matches 34; Conservative 21; Mismatches 50; Indels 36; Gaps 6

QY      61 NATLLYNNKHLTDIKESHETITVKKAYDYTTENGIKY--ETQGR-----TQTFEDVYV 112
      1 : 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      289 NSTTPYCNLDLKKKNLESSPLSRITWKALKPLLVGKILYPPDPATPQVAEYVKKTFQELAY 348
      1 : 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      113 FSDYKNCDDVIVPK-----ERGSDEGELYELWVSDDKIDKIDCKKF 153
      1 : 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      349 PHDEEGMEELSPQITWTFMENSQEMDLVRLTDSGRND---QEW--EQKIDGLDWTAQD 402
      1 : 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      154 TMAVFAQOOEKTVR--NVYT 171
      1 : 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      403 IMAFLAKNPEDVQSPNGSVYT 423
      1 : 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-09-893-737-172
; Sequence 172, Application US/09893737
; Patent No. US2002011085A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41

```

```

:
: CURRENT APPLICATION NUMBER: US/09/893,737
:
: CURRENT FILING DATE: 2001-06-28
:
: PRIOR APPLICATION NUMBER: US 60/215,446
:
: PRIOR FILING DATE: 2000-06-30
:
: NUMBER OF SEQ ID NOS: 329
:
: SOFTWARE: FASTSEQ for Windows Version 3.0.
:
: SEQ ID NO 172
:
: LENGTH: 142
:
: TYPE: prt
:
: ORGANISM: Homo sapiens
:
: US-09-893-737-172

```

Query Match	7.5%	Score 74.5;	DB 10;	Length 142;
Best Local Similarity	20.8%	Pred. No. 2.7;		
Matches	33;	Conservative	28;	Mismatches 51;
				Indels 47;
				Gaps 7

```

QY 30 LAATQRTQTSBWMEERTCSVTAE--KIGKK-----LNMILYKXNHLDTLES--HET 80
    | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 15 LAEASEETMKIIEETDEEMQYDAKSQEQOTIEILNPIIVKNTLSMSKDDMSST 74
    | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY 81 ITVWKAADVTTENGIVETFGTRTQTFEDYVFPFSYKNCQDVIYVPKKSGSDE--GPYEL 138
    | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 75 LITFRSLHYIDPKG-----NNSGNDKECCNDMTV 103
    | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY 138 WVSQDKIDKIPDCCKFTYMAFYAQOQDEKTVRNVYTTDSCK 176
    | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 104 W--RRKYSEANGSCCKWS--NNFTRSSTEVARRRHAPSPCK 138
    | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

RESULT 11  
 US-09-864-761-48672  
 : Sequence 48672, Application US/09864761  
 : Patent No. US20020048763A1  
 :  
 : GENERAL INFORMATION:  
 : APPLICANT: Penn, Sharon G.  
 : APPLICANT: Rank, David R.  
 : APPLICANT: Hanzel, David K.  
 : APPLICANT: Chen, Wensheng  
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 : TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 : FILE REFERENCE: Acomica-x-1  
 : CURRENT APPLICATION NUMBER: US/09/864,761  
 : CURRENT FILING DATE: 2001-05-23  
 : PRIOR APPLICATION NUMBER: US 60/180,312  
 : PRIOR FILING DATE: 2000-02-04  
 : PRIOR APPLICATION NUMBER: US 60/207,456  
 : PRIOR FILING DATE: 2000-05-26  
 : PRIOR APPLICATION NUMBER: US 09/632,366  
 : PRIOR FILING DATE: 2000-08-03  
 : PRIOR APPLICATION NUMBER: GB 24263,6  
 : PRIOR FILING DATE: 2000-10-04  
 : PRIOR APPLICATION NUMBER: US 60/236,359  
 : PRIOR FILING DATE: 2000-09-27  
 : PRIOR APPLICATION NUMBER: PCT/US01/00666  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00667  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00664  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00669  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00665  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00668  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00663  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00662  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00661  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00670  
 : PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48672
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010766.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: AW188146.1, EVALU 1.00e-77
; OTHER INFORMATION: SWISSPROT HIT: P00673, EVALU 1.00e-05
US-09-864-761-48672
```

```
Query Match
Best Local Similarity 7.5%; Score 74.5; DB 9; Length 198;
Matches 33; Conservative 28; Mismatches 51; Indels 47; Gaps 7;
```

```
OY 30 LAQATQTTDGVGGEFTCVSYAE--KIGKK-----LNATILYKKNHLDLKS--HET 80
DB 14 LAEASESTMIKEEFTDEEMQYMAKSGQKQTEILMNPILLVKNLSLSKSDMST 73
OY 81 ITWKAYDYTTENGIKETGCTGTOTFEDEVFVSDYKNCVDYFVPKKGSDP---GDYEL 137
DB 74 LITFRSLHYNDPKG-----NSSGNDKECCNDMTV 102
OY 138 WVEDKIDKIDPCCKFTMAVFAOQOEKTVRVNYTDSCK 176
DB 103 W---RKVSEANGSKMS--NFIIRSTEYMRVHRAPSK 137
```

```
RESULT 12
US-09-801-231-2
; Sequence 2, Application US/09801231
; Patent No. US20010049434A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: Human Ribonuclease
; FILE REFERENCE: 00-19
; CURRENT FILING DATE: 2001-03-07
; CURRENT APPLICATION NUMBER: US/09/801,231
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-801-231-2
```

```
Query Match
Best Local Similarity 7.5%; Score 74.5; DB 9; Length 199;
Matches 33; Conservative 28; Mismatches 51; Indels 47; Gaps 7;
```

```
OY 30 LAQATQTTDGVGGEFTCVSYAE--KIGKK-----LNATILYKKNHLDLKS--HET 80
DB 15 LAEASESTMIKEEFTDEEMQYMAKSGQKQTEILMNPILLVKNLSLSKSDMST 74
OY 81 ITWKAYDYTTENGIKETGCTGTOTFEDEVFVSDYKNCVDYFVPKKGSDP---GDYEL 137
DB 75 LITFRSLHYNDPKG-----NSSGNDKECCNDMTV 103
OY 138 WVEDKIDKIDPCCKFTMAVFAOQOEKTVRVNYTDSCK 176
DB 104 W---RKVSEANGSKMS--NFIIRSTEYMRVHRAPSK 137
```

```
RESULT 13
US-09-731-872-269
```

```
; Sequence 269, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouguetelert, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78-US3-REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 269
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -16...-1
US-09-731-872-269
```

```
Query Match
Best Local Similarity 7.5%; Score 74.5; DB 10; Length 199;
Matches 33; Conservative 28; Mismatches 51; Indels 47; Gaps 7;
```

```
OY 30 LAQATQTTDGVGGEFTCVSYAE--KIGKK-----LNATILYKKNHLDLKS--HET 80
DB 15 LAEASESTMIKEEFTDEEMQYMAKSGQKQTEILMNPILLVKNLSLSKSDMST 74
OY 81 ITWKAYDYTTENGIKETGCTGTOTFEDEVFVSDYKNCVDYFVPKKGSDP---GDYEL 137
DB 75 LITFRSLHYNDPKG-----NSSGNDKECCNDMTV 103
OY 138 WVEDKIDKIDPCCKFTMAVFAOQOEKTVRVNYTDSCK 176
DB 104 W---RKVSEANGSKMS--NFIIRSTEYMRVHRAPSK 137
```

```
RESULT 14
US-10-052-586-588
; Sequence 588, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT FILING DATE: 2002-01-15
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
```

1	PRIOR APPLICATION NUMBER: 60/0634868
2	PRIOR FILING DATE: 1997-10-21
3	PRIOR APPLICATION NUMBER: 60/0635400
4	PRIOR FILING DATE: 1997-10-28
5	PRIOR APPLICATION NUMBER: 60/0635411
6	PRIOR FILING DATE: 1997-10-28
7	PRIOR APPLICATION NUMBER: 60/0635444
8	PRIOR FILING DATE: 1997-10-28
9	PRIOR APPLICATION NUMBER: 60/0635544
10	PRIOR FILING DATE: 1997-10-28
11	PRIOR APPLICATION NUMBER: 60/0635544
12	PRIOR FILING DATE: 1997-10-28
13	PRIOR APPLICATION NUMBER: 60/0637344
14	PRIOR FILING DATE: 1997-10-29
15	PRIOR APPLICATION NUMBER: 60/0638700
16	PRIOR FILING DATE: 1997-10-31
17	PRIOR APPLICATION NUMBER: 60/0641030
18	PRIOR FILING DATE: 1997-10-31
19	PRIOR APPLICATION NUMBER: 60/0653111
20	PRIOR FILING DATE: 1997-11-13
21	PRIOR APPLICATION NUMBER: 60/0661200
22	PRIOR FILING DATE: 1997-11-21
23	PRIOR APPLICATION NUMBER: 60/0664666
24	PRIOR FILING DATE: 1997-11-24
25	PRIOR APPLICATION NUMBER: 60/0667722
26	PRIOR FILING DATE: 1997-11-24
27	PRIOR APPLICATION NUMBER: 60/0693355
28	PRIOR FILING DATE: 1997-12-11
29	PRIOR APPLICATION NUMBER: 60/0694255
30	PRIOR FILING DATE: 1997-12-12
31	PRIOR APPLICATION NUMBER: 60/0698700
32	PRIOR FILING DATE: 1997-12-17
33	PRIOR APPLICATION NUMBER: 60/0680177
34	PRIOR FILING DATE: 1997-12-18
35	PRIOR APPLICATION NUMBER: 60/0774500
36	PRIOR FILING DATE: 1998-03-10
37	PRIOR APPLICATION NUMBER: 60/0776322
38	PRIOR FILING DATE: 1998-03-11
39	PRIOR APPLICATION NUMBER: 60/0776649
40	PRIOR FILING DATE: 1998-03-11
41	PRIOR APPLICATION NUMBER: 60/0788866
42	PRIOR FILING DATE: 1998-03-20
43	PRIOR APPLICATION NUMBER: 60/0789399
44	PRIOR FILING DATE: 1998-03-20
45	PRIOR APPLICATION NUMBER: 60/0796644
46	PRIOR FILING DATE: 1998-03-27
47	PRIOR APPLICATION NUMBER: 60/0797866
48	PRIOR FILING DATE: 1998-03-27
49	PRIOR APPLICATION NUMBER: 60/0801077
50	PRIOR FILING DATE: 1998-03-31
51	PRIOR APPLICATION NUMBER: 60/0801944
52	PRIOR FILING DATE: 1998-03-31
53	PRIOR APPLICATION NUMBER: 60/0803327
54	PRIOR FILING DATE: 1998-04-01
55	PRIOR APPLICATION NUMBER: 60/0810499
56	PRIOR FILING DATE: 1998-04-08
57	PRIOR APPLICATION NUMBER: 60/0810700
58	PRIOR FILING DATE: 1998-04-08
59	PRIOR APPLICATION NUMBER: 60/0811955
60	PRIOR FILING DATE: 1998-04-09
61	PRIOR APPLICATION NUMBER: 60/0818388
62	PRIOR FILING DATE: 1998-04-15
63	PRIOR APPLICATION NUMBER: 60/0825568
64	PRIOR FILING DATE: 1998-04-21
65	PRIOR APPLICATION NUMBER: 60/0825699
66	PRIOR FILING DATE: 1998-04-21
67	PRIOR APPLICATION NUMBER: 60/0827004
68	PRIOR FILING DATE: 1998-04-22
69	PRIOR APPLICATION NUMBER: 60/0827977
70	PRIOR FILING DATE: 1998-04-22
71	PRIOR APPLICATION NUMBER: 60/0833222
72	PRIOR FILING DATE: 1998-04-28
73	PRIOR APPLICATION NUMBER: 60/0834955

PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083496
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083495
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083557
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/084366
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084634
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/084414
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION NUMBER:	60/084633
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/085577
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085579
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085580
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085582
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085700
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/086023
PRIOR FILING DATE:	1998-05-18
PRIOR APPLICATION NUMBER:	60/086392
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/086486
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/087098
PRIOR FILING DATE:	1998-05-28
PRIOR APPLICATION NUMBER:	60/087208
PRIOR FILING DATE:	1998-05-28
PRIOR APPLICATION NUMBER:	60/087609
PRIOR FILING DATE:	1998-06-02
PRIOR APPLICATION NUMBER:	60/087759
PRIOR FILING DATE:	1998-06-02
PRIOR APPLICATION NUMBER:	60/088029
PRIOR FILING DATE:	1998-06-03
PRIOR APPLICATION NUMBER:	60/088025
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088028
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088029
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088033
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088167
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/088202
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/088212
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/088217
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/088326
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088655
PRIOR FILING DATE:	1998-06-09
PRIOR APPLICATION NUMBER:	60/088722
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088738
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088740
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088811
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088824
PRIOR FILING DATE:	1998-06-10





THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 13.8667 Seconds

(Without alignments)  
555.330 Million cell updates/sec

Title: US-10-087-195-8

Perfect score: 989

Sequence: 1 NPTWANEAKGSGYQDAWKSL.....EKTIVRNYTDSCKPAPAQN 182

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/Backfilled.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79.5	8.0	202	1 US-07-807-022A-1	Sequence 1, Appli
2	78	7.9	1095	4 US-09-107-532A-3855	Sequence 3855, Ap
3	77.5	7.8	468	4 US-09-107-532A-5304	Sequence 5304, Ap
4	77.5	7.8	823	4 US-09-107-532A-5667	Sequence 5667, Ap
5	77	7.8	341	4 US-09-563-269-19	Sequence 19, Appl
6	76.5	7.7	216	3 US-08-824-692-24	Sequence 24, Appl
7	75.5	7.6	347	4 US-03-328-352-6764	Sequence 6764, Ap
8	75	7.6	500	4 US-09-071-035-396	Sequence 396, App
9	75	7.6	1074	4 US-09-071-035-358	Sequence 358, App
10	75	7.6	1074	4 US-09-071-035-394	Sequence 394, App
11	74.5	7.5	334	4 US-09-198-452A-1120	Sequence 1120, App
12	74	7.5	771	4 US-09-634-238-329	Sequence 229, App
13	74	7.5	771	4 US-09-634-238-262	Sequence 262, App
14	73.5	7.4	144	4 US-09-253-991A-23258	Sequence 23258, A
15	73	7.4	325	3 US-08-961-083-80	Sequence 80, Appl
16	73	7.4	325	4 US-09-536-784-80	Sequence 80, Appl
17	72.5	7.3	401	4 US-09-253-991A-17272	Sequence 17272, A
18	72.5	7.3	565	4 US-08-961-083-218	Sequence 218, App
19	72.5	7.3	565	4 US-09-536-784-218	Sequence 218, App
20	72.5	7.3	901	2 US-08-884-681-5	Sequence 5, Appl1
21	72.5	7.3	901	2 US-09-258-643-5	Sequence 5, Appl1
22	72	7.3	1244	5 PCT-US93-10500-2	Sequence 2, Appl1
23	71.5	7.2	397	1 US-08-290-448A-76	Sequence 76, Appl
24	71.5	7.2	397	1 US-08-290-448A-76	Sequence 76, Appl
25	71.5	7.2	397	1 US-08-175-069A-76	Sequence 76, Appl
26	71.5	7.2	397	1 US-08-461-939B-76	Sequence 76, Appl
27	71.5	7.2	397	4 US-08-464-000-76	Sequence 76, Appl

28	71.5	7.2	805	4 US-09-134-001C-4821	Sequence 4821, Ap
29	71	7.2	620	4 US-08-637-670-40	Sequence 40, Appl
30	71	7.2	648	4 US-09-199-637A-221	Sequence 221, App
31	71	7.2	972	3 US-08-335-844A-24	Sequence 24, Appl
32	71	7.2	972	4 US-09-129-366-24	Sequence 23, Appl
33	70.5	7.1	240	3 US-08-824-692-23	Sequence 23, Appl
34	70.5	7.1	246	4 US-09-107-532A-3862	Sequence 3862, Ap
35	70	7.1	816	3 US-09-113-750A-34	Sequence 34, Appl
36	70	7.1	471	4 US-09-134-001C-4125	Sequence 4125, Ap
37	70	7.1	603	4 US-09-508-264A-1	Sequence 1, Appl1
38	70	7.1	845	4 US-08-851-567B-37	Sequence 37, Appl
39	70	7.1	851	4 US-09-071-035-326	Sequence 326, App
40	70	7.1	851	4 US-09-071-035-330	Sequence 330, App
41	70	7.1	851	4 US-09-071-035-334	Sequence 334, App
42	70	7.1	862	4 US-09-206-551-15	Sequence 15, Appl
43	70	7.1	890	1 US-08-472-934-8	Sequence 8, Appl1
44	70	7.1	890	2 US-08-323-460A-8	Sequence 8, Appl1
45	70	7.1	890	2 US-08-461-146C-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-07-807-022A-1  
Sequence 1, Application US/07807022A  
Patent No. 5321010

## GENERAL INFORMATION:

APPLICANT: Connolly, Thomas M.

APPLICANT: Neepert, Michael

APPLICANT: Waxman, Lloyd

TITLE OF INVENTION: Protein for Inhibiting Adhesion of

TITLE OF INVENTION: Platelets to Collagen

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Merck &amp; Co., Inc.

STREET: P.O. Box 2000

CITY: Rahway

STATE: New Jersey

COUNTRY: USA

ZIP: 08840

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/807,022A

FILING DATE: 19911210

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Parr, Richard S.

REGISTRATION NUMBER: 32,586

REFERENCE/DOCKET NUMBER: 18143

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-4958

TELEFAX: 908-594-4720

TELEX: 138825

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 202 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-07-807-022A-1

Query Match 8.0%; Score 79.5; DB 1; Length 202;  
Best Local Similarity 21.4%; Pred. No. 0.75; Indels 27; Gaps 8;  
Matches 39; Conservative 30; Mismatches 86;  
14 ODANKSLQDDNKRRYLAQATQTTDGCVGEEFTCVSAEKIGKKLNATLLYKNNKHLTD 73

Db 33 RDIKKT--YKSKYFVFLIRTYWD--GRSLCRYGTVLK--RDKANHRI---EQING 82  
Qy 74 LKESHEITVWKAYDYT-----TENGIKYETQTRQTFEDVFEVSDKNCQDYFVPE 127  
Db 83 YDESEYETKTKVYITTRILGRSGDRNHGVSLOQYVNHGIEYKMIYDDOCALIKYTKD 142  
Qy 128 RGSQ-----EGDYELWVSEDKIDKIPDOCKFTMAVFAOQOEXTVRNVYTD---SSCPAP 179  
Db 143 NRNPONLKNACEMWATANDANSVSIACEVY---ORCNPNNSVDPPYVTKWMP 199  
Qy 180 AQ 181  
Db 200 VE 201  
RESULT 2  
US-09-107-532A-3855  
Sequence 3855, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3855:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1095 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...1095  
SEQUENCE DESCRIPTION: SEQ ID NO: 3855:  
US-09-107-532A-3855  
Query Match 7.98; Score 78; DB 4; Length 1095;  
Best Local Similarity 23.08; Pred. No. 12;  
Matches 49; Conservative 34; Mismatches 64; Indels 66; Gaps 13;  
Qy 15 DANKSLQD--DUNKRYLAOTOT-----TDGVWGEFTCVSYAEKIGKKLNAT--- 64  
Db 819 DAKAOKDEKAKPVYIAETSATLANKETGTWKIOH---KLPAQV---LKKTVLF 871

Qy 65 --LYKNKH-----LTDKESHEITVWKAYDYTTENGIKYETQ-----GTRTOTFE 108  
Db 872 NYVENKPAFAGDKPAKADVSLNNQAQTV-----SCYEHVSVIQTANLHNSQITTHG 927  
Qy 109 DVF-VFSDYK-NCQVIVPKERGS-----DEGDYELW-----VSE 141  
Db 928 DVMFDFDVSITHHVLSGSKAEFTIYALPDDTNRKIMWSKIEYEVDNKEFTKYVLA 987  
Qy 142 DKID--KIPDOCKFTMAVFAOQOEXTVRNVYTD 172  
Db 988 KVDTKGKPEGTGKTFPAINDKDGTINGKNE 1020

RESULT 3  
US-09-107-532A-5304  
Sequence 5304, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5304:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 468 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...468  
SEQUENCE DESCRIPTION: SEQ ID NO: 5304:  
US-09-107-532A-5304  
Query Match 7.88; Score 77.5; DB 4; Length 468;  
Best Local Similarity 24.18; Pred. No. 4.1;  
Matches 32; Conservative 16; Mismatches 50; Indels 35; Gaps 4;  
Qy 12 SYDANKSLQDQDNKRYLAOTOTTDGVWGEFTCVSYAEKIGKKLNA-TILYKNKH 70  
Db 12 NYPLWPRPKOKLKRPIYRSLSQD-----LEQDILKNLQKNTQLPSQRE 56

```

OY 71 TLDKESHEFTVAKADYTTENGIKETQSTRQTEDEVF-----V 112
Db 57 LADLDLNET-TVGAKAKYGIKGLTYNIGSGTFPSNPFESTISTNEVTEHIIDFGL 115
OY 113 FSDYKNCDFVFP 125
Db 116 VSSFECNKFVP 128

RESULT 4
US-09-107-532A-5667
: Sequence 5667, Application US/09107532A
: Patent No. 6583275
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7310
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER: PC
: OPERATING SYSTEM: <Unknown>
: SOFTWARE: ASCII
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,532A
: FILING DATE: 30-Jun-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/085,598
: FILING DATE: 14 May 1998
: APPLICATION NUMBER: 60/051571
: FILING DATE: July 2, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Arinello, Pamela Deneke
: REGISTRATION NUMBER: 40,489
: REFERENCE/DOCKET NUMBER: GTC-012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781)993-5007
: TELEFAX: (781)893-8277
: INFORMATION FOR SEQ ID NO: 5667:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 823 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEetical: YES
: ORIGINAL SOURCE:
: ORGANISM: Enterococcus faecium
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (B) LOCATION 1...823
: SEQUENCE DESCRIPTION: SEQ ID NO: 5667:
US-09-107-532A-5667

Query Match 7.88; Score 77.5; DB 4; Length 823;
Best Local Similarity 24.38; Pred. No. 9.1;
Matches 42; Conservative 28; Mismatches 62; Indels 41; Gaps 10;

OY 14 ODAKKSLOODNKR-----YLLAOTATOTDGSWMGEFFCVSVTAEKIGKKKLTILYKN 68
Db 210 ODAAMAAVRLKKSKOELLITYVKNKYMSNGLGHE-----TASEMFGKKLSLSLPOT 264
OY 69 KHLTDLKESHTTIVWKAVD-YTTENGIKIYETOGTRTQTEFEDVFVSDYKNCADV----- 121
Db 265 ALLAGMGQAP-----SAYDPIY-----YEDQAKRR-----DYLVLYTMQLQNEKXSQTFEYD 309
OY 122 --IFVPERGSDCEDDYELWVSEDKIKIDPCCKFTMAVFAOQOEKTVARNVYTD 172

```

```

Db      : 310 QAVANP-----VTDLQELTGSDDN-TKIVD--NYVKEYINEVOETKKNVITD 355

RESULT 5
US-09-563-269-19
; Sequence 19, Application US/09563269
; Patent No. 6553655
; GENERAL INFORMATION:
; APPLICANT: RUDAR, MARK J.
; APPLICANT: DONOVAN, WILLIAM P.
; APPLICANT: CHU, CHI-HREI
; APPLICANT: PEASE, ELIZABETH
; APPLICANT: TAN, YUPING
; APPLICANT: SLANEY, ANNETTE C.
; APPLICANT: BAUM, JAMES A.
; APPLICANT: MALVAR, THOMAS M.
; TITLE OF INVENTION: COLOSTERAN-TOXIC POLYPEPTIDE COMPOSITIONS AND INSECT
; TITLE OF INVENTION: RESISTANT TRANSGENIC PLANTS
; FILE REFERENCE: MECO164
; CURRENT APPLICATION NUMBER: US/09/563,269
; CURRENT FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-09-563-269-19

Query Match          7.8%; Score 77; DB 4; Length 341:
Best Local Similarity 24.3%; Pred. No. 3;
Matches 28; Conservative 26; Mismatches 49; Indels 12; Gaps 4

Qy      22 ODNKKRYLAATOTDGCWGEBCFVCVTAEKIKGKKLNATILYKNKHLDTKESHETI 81
       1 : :: | | | : : | : | : | : | : | : | : | : | : | : | : | : | :
Db      213 OEKSORNLALTDYSKEQACKKIT-MNVITADLTAKTPG----YNPPRRYEDDSHNNL 267
       1 : : : | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy      82 TWVK-----AYDTTENGIRYETOGTRTOTFEDEVFVS DYKNCDFIEVPKERGS 130
       1 : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      268 TTHSDLIYNLEDFAFYEITVFENLSYSTLFNFEDLFYRFDKNNHL-LIEKT VGS 321
       1 : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 6
US-08-824-692-24
; Sequence 24, Application US/08824692
; Patent No. 6017703
; GENERAL INFORMATION:
; APPLICANT: Kinders, Robert J.
; APPLICANT: Enfield, David L.
; APPLICANT: Hass, G. Michael
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
; TITLE OF INVENTION: FOR OR MODULATING A TUMOR ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,692
; FILING DATE: 08-Apr-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629

```

REFERENCE/DOCKET NUMBER: 130001.404  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 216 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-824-692-24

Query Match 7.7%; Score 76.5; DB 3; Length 216;  
Best Local Similarity 19.5%; Pred. No. 1.8;  
Matches 36; Conservative 31; Mismatches 69; Indels 49; Gaps 9;

QY 19 SLQDDQKKRYLAQATQ-----TTGCGWGEER-TCVSTAEK---IGK 57  
DB 47 SSAMEPDEHYHFGAVFVCSGYKIEGDEEMHCSDDGFGWKEKPKCWEISCKSPDYING 106  
QY 58 KLNATILYKNNKHLTDLKESHETITWKADYTTENGICYETQGT--TQTFEYFVPSD 115  
DB 107 SPIQKIIYKEN-----ERFOYKCNMGYEYSEKDAVCTESGMRPLPSC 151  
QY 116 YKNCQVITVPKERSGSDGDIY-LWVSEDKIDKIPDCKFTWAFPAQOQEKTVRNVTSS 174  
DB 152 EKSCDNPYIP-----NGDYSPLIKHRTGDEITYQCR--NGFYPATRGNTAK--CSTG 201  
QY 175 CKRAP 179  
DB 202 WIRAP 206

RESULT 7  
US-09-328-352-6764  
Sequence 6764, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328.352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 6764  
LENGTH: 347  
TYPE: PRP  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-6764

Query Match 7.6%; Score 75.5; DB 4; Length 347;  
Best Local Similarity 22.2%; Pred. No. 4.5;  
Matches 24; Conservative 19; Mismatches 42; Indels 23; Gaps 4;

QY 6 NEAKLSYQDAWKSLOO--DQNKRYLAQATQTTDGVWGEETCVSTAEKIKKKLNAT 63  
DB 234 NMSHLTAHQEEMKKIIQWMDKTYKINDPATHA-----DAVKIMARSGVDPRQYE 264  
QY 64 ILKNNKHLTDLKESHETITWKAYD-----YTTENGICYETQ 100  
DB 285 LMSVSGTHLDINANKKVFYKSGFSDSYSHVNNFENVENGI-YKTE 331

RESULT 8  
US-09-071-035-396  
Sequence 396, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 396:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 500 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-396

Query Match 7.6%; Score 75; DB 4; Length 500;  
Best Local Similarity 28.2%; Pred. No. 8.5;  
Matches 22; Conservative 13; Mismatches 29; Indels 14; Gaps 3;

QY 11 GSYQDAWKS--LQDDQKKRY---LAQATQTTDGVWGEETCVSTAEKIKKKLNAT 64  
DB 400 GINKEIKWSKIEHVNDEKFTYVLAEKYDTGKYPBGIKFTTETINEYKDN----- 452  
QY 65 LYRKNHLTDLKESHETIT 82  
DB 453 -VNGKHNEDEKESQTLF 469

RESULT 9  
US-09-071-035-358  
Sequence 358, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

```

; TYPE: PRI
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-229

```

Query Match 7.58; Score 74; DB 4; Length 771;  
Best Local Similarity 23.98; Pred. No. 20;  
Matches 28; Conservative 16; Mismatches 47; Indels 26; Gaps 4;

QY 29 YLAQATQTTDGVWGEFTCVSVTAERKIGKKLNATILYKKNHLDLTKESHETITWKAYD 88  
DB 571 YQAGKTGTTD--YSDE-----ELKONPALNATGIAKAMFTGYR--NRVTSVWTGYD 619  
QY 89 YTTENGKIKYETGCTRTQTFEDVFVSDYKNCDDVIFVPERKSGDEGYELWVSEDKID 145  
DB 620 KPTSHGISYAEQRTISQIKYKALMSYT-----SONLDNKDWTKPPDPTVE 661

RESULT 13  
US-09-634-238-262  
; Sequence 262, Application US/09634238  
; Patent No. 6544772  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Havukkala, Ilkka J.  
; APPLICANT: Bloksberg, Leonard, N.  
; APPLICANT: Lubbers, Mark W.  
; APPLICANT: Dekker, James  
; APPLICANT: Christensson, Anna C.  
; APPLICANT: Holland, Ross  
; APPLICANT: O'Toole, Paul W.  
; APPLICANT: Reid, Julian R.  
; APPLICANT: Coolbear, Timothy  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; FILE REFERENCE: 11000.104301  
; CURRENT APPLICATION NUMBER: US/09/634.238  
; CURRENT FILING DATE: 2000-08-08  
; NUMBER OF SEQ ID NOS: 422  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 262  
; LENGTH: 771  
; TYPE: PRT  
; ORGANISM: Lactobacillus rhamnosus  
; US-09-634-238-262

Query Match 7.58; Score 74; DB 4; Length 771;  
Best Local Similarity 23.98; Pred. No. 20;  
Matches 28; Conservative 16; Mismatches 47; Indels 26; Gaps 4;

QY 29 YLAQATQTTDGVWGEFTCVSVTAERKIGKKLNATILYKKNHLDLTKESHETITWKAYD 88  
DB 571 YQAGKTGTTD--YSDE-----ELKONPALNATGIAKAMFTGYR--NRVTSVWTGYD 619  
QY 89 YTTENGKIKYETGCTRTQTFEDVFVSDYKNCDDVIFVPERKSGDEGYELWVSEDKID 145  
DB 620 KPTSHGISYAEQRTISQIKYKALMSYT-----SONLDNKDWTKPPDPTVE 661

RESULT 14  
US-09-252-991A-23258  
; Sequence 23258, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23258  
; LENGTH: 144  
; TYPE: PRT

ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23258

Query Match 7.48; Score 73.5; DB 4; Length 144;  
Best Local Similarity 26.78; Pred. No. 2.1;  
Matches 27; Conservative 14; Mismatches 41; Indels 19; Gaps 4;

QY 5 ANEAKLSYODAMKSLQODQ--NKRYYLAQATQTTDGVWGEFTCVSVTAER--KIGK-- 57  
DB 29 ATEFAIKELQERLKSLEQDDKLKKELEFEKRLTLMGTQKSLRDVLSLDPDAKIGKST 88  
QY 58 -----KKLNATILYKKNHLDLKE-----SHETITWKA 86  
DB 89 RTAKAPAKRRARVKYKNPHTGEVETKGNHKTLEWKA 129

RESULT 15  
US-08-961-083-80  
; Sequence 80, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 80:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 325 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-961-083-80

Query Match 7.48; Score 73; DB 3; Length 325;  
Best Local Similarity 19.58; Pred. No. 7.7;  
Matches 40; Conservative 31; Mismatches 84; Indels 50; Gaps 8;

QY 13 YODA-NKSLQODQNKRRYYLAQATQTTDGVWGEFTCVSVTAER-----IGKK--LN 61  
DB 96 YDAPMTPNKDGNGNLTEGYFLQRLNEVAKLEIPDVMAIAESSSAIKITGKEIGGLG 155  
QY 62 ATILYKKNHLDLTKESHETITWKAYDTT-----ENGK 96  
DB 156 FDYKMNKGMNDLRLRYEDPIRYKDFNLFVTFSPVYKKNENLLPFSHDEVVHGKSKM 215  
QY 97 YETGCTRTQTFEDVFVSDYKNC-----DVIFVPERKSGDEGYELWVSEDKID--KIPDC 150



Db	216	HKMWDRYNOFAGLRNLVYQICHGKKLFFM---- <td>271</td>	271
Oy	151	CKFTMAYFAOQOEKTVRNVTDSGC	175
Db	272	MNAKMKTFASO----LNQFYKDHRG	292

Search completed: August 1, 2003, 12:17:22  
Job time : 15.8667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 56.5067 Seconds  
(without alignments)  
511.236 Million cell updates/sec

Title: US-10-087-195-8  
Perfect score: 989  
Sequence: 1 NPTWANEAKLGSYODAMKSL.....EKTVRNVTYDSCCKPAPARN 182

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A.GeneSeq.19Jun03.\*  
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.\*  
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.\*  
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.\*  
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.\*  
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.\*  
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.\*  
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.\*  
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.\*  
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.\*  
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.\*  
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.\*  
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.\*  
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.\*  
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.\*  
25: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	989	100.0	182	AA1980	Histacalin protein
2	989	100.0	182	AA1981	Histacalin protein
3	989	100.0	200	AA1982	Tick vasocative am
4	989	100.0	200	AA1983	Histamine binding
5	331.5	33.5	171	AA1984	Histamine binding
6	331.5	33.5	171	AA1985	Histamine binding
7	331.5	33.5	190	AA1986	Histamine binding
8	331.5	33.5	190	AA1987	Histamine binding
9	267.5	27.0	172	AA1988	Histacalin protein

10	267.5	27.0	172	AA1989	Histacalin protein
11	267.5	27.0	190	AA1990	Tick vasocative am
12	267.5	27.0	190	AA1991	Histamine binding
13	267.5	27.0	203	AA1992	Histamine binding
14	263	26.6	203	AA1993	Histamine binding
15	251.5	25.4	209	AA1994	Histamine binding
16	251.5	25.4	209	AA1995	Histamine binding
17	243	24.6	198	AA1996	Histamine binding
18	123.5	12.5	285	AA1997	Histamine binding
19	119.5	12.1	284	AA1998	Histamine binding
20	109.5	11.1	161	AA1999	Human ORP protein
21	101.5	10.3	207	AA2000	Histamine binding
22	92.5	9.4	321	AA2001	Histamine binding
23	87	8.8	483	AA2002	Listeria monocytog
24	85	8.6	514	AA2003	Theobrom cacao as
25	85	8.6	1959	AA2004	Listeria monocytog
26	83.5	8.4	778	AA2005	Enterococcus faeca
27	82.5	8.3	1802	AA2006	Plasmodium falcipa
28	82	8.3	392	AA2007	Streptococcus poly
29	80.5	8.1	721	AA2008	Listeria monocytog
30	79.5	8.0	202	AA2009	Sequence of prote
31	79.5	8.0	2681	AA2010	Pathogen specific
32	79	8.0	221	AA2011	I. scapularis salp
33	79	8.0	375	AA2012	Listeria monocytog
34	79	8.0	396	AA2013	Ragweed allergen 1
35	79	8.0	431	AA2014	Human secreted pro
36	79	8.0	541	AA2015	Human secreted pro
37	79	8.0	542	AA2016	Human secreted pro
38	79	8.0	1491	AA2017	Human secreted pro
39	79	8.0	1605	AA2018	Novel human secret
40	79	8.0	1607	AA2019	Human mitogen-acti
41	79	8.0	1830	AA2020	Urea amide lyase.
42	78.5	7.9	429	AA2021	Candida albicans e
43	77.5	7.8	423	AA2022	Listeria monocytog
44	77	7.8	341	AA2023	Bacillus thuringie
45	77	7.8	745	AA2024	Human secretory po

ALIGNMENTS

RESULT 1  
ID AAB74290 standard; protein: 182 AA.  
AC AAB74290:  
20-JUN-2001 (first entry)  
Histacalin protein MS-HBPL.  
Histacalin: MS-HBPL; conjunctivitis.  
Unidentified.  
WO200115719-A2.  
08-MAR-2001.  
24-AUG-2000; 2000WO-GB03282.  
01-SEP-1999; 99GB-0020674.  
(EVOL-) EVOLUTEC LTD.  
Nutall PA, Paesen GC;  
WPI: 2001-257675/26.  
Use of histacalin proteins for treating or preventing non-infective  
conjunctivitis, or for manufacturing a medicament for treating or  
preventing conjunctivitis, e.g. seasonal or perennial allergic  
conjunctivitis

XX Claim 4; Page 5-6; 19pp; English.  
PS  
XX The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein PS-HBPI. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.  
XX  
SQ Sequence 182 AA;

Query Match 100.0%; Score 989; DB 22; Length 182;  
Best Local Similarity 100.0%; Pred. No. 3,1e-94;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPTWANEAKLGSYDADKSLQODONKRYRYLAQATQTTDYGWGEFFTCVSTAERIGKKKL 60  
DB 1 NPTWANEAKLGSYDADKSLQODONKRYRYLAQATQTTDYGWGEFFTCVSTAERIGKKKL 60  
QY 61 NATLYNKKHLTDLKESHETITWKAYDYTTENGIKYETGTRQTFEDVFEVSDYKNC 120  
DB 61 NATLYNKKHLTDLKESHETITWKAYDYTTENGIKYETGTRQTFEDVFEVSDYKNC 120  
QY 121 VIFVPEKRGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAQQOKEKTVANNVTTDSSCKRPAPA 180  
DB 121 VIFVPEKRGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAQQOKEKTVANNVTTDSSCKRPAPA 180  
QY 181 QN 182  
DB 181 QN 182

## RESULT 2

ID AAB73262 standard; protein; 182 AA.

AC AAB73262;

DT 15-MAY-2001 (first entry)

DE Histacalin protein MS-HBPI.

XX MS-HBPI; histacalin protein; antinflammatory; antiallergic;

KW ophthalmological; allergic rhinitis.

XX Unidentified.

OS WO200116164-A2.

PN 08-MAR-2001.

PD 24-AUG-2000; 2000WO-GB03287.

PE 01-SEP-1999; 99GB-0020673.

PR (EVOL-) EVOLUTEC LTD.

PA Nuttall PA, Paesen GC;

XX WPI; 2001-218521/22.

DR Use of histacalin proteins for treating or preventing allergic

XX rhinitis, or for manufacturing a medicament for treating or preventing

PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis

XX Disclosure; Pages 4-6; 19pp; English.

XX The present invention relates to a method for treating or preventing  
CC allergic rhinitis. The method involves employing a blood-feeding  
CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
CC sequence is one such histacalin protein. The histacalin protein, is  
CC useful for treating or preventing allergic rhinitis, both seasonal and  
CC perennial allergic conjunctivitis.

SQ Sequence 182 AA;

Query Match 100.0%; Score 989; DB 22; Length 182;  
Best Local Similarity 100.0%; Pred. No. 3,1e-94;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPTWANEAKLGSYDADKSLQODONKRYRYLAQATQTTDYGWGEFFTCVSTAERIGKKKL 60  
DB 1 NPTWANEAKLGSYDADKSLQODONKRYRYLAQATQTTDYGWGEFFTCVSTAERIGKKKL 60  
QY 61 NATLYNKKHLTDLKESHETITWKAYDYTTENGIKYETGTRQTFEDVFEVSDYKNC 120  
DB 61 NATLYNKKHLTDLKESHETITWKAYDYTTENGIKYETGTRQTFEDVFEVSDYKNC 120  
QY 121 VIFVPEKRGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAQQOKEKTVANNVTTDSSCKRPAPA 180  
DB 121 VIFVPEKRGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAQQOKEKTVANNVTTDSSCKRPAPA 180  
QY 181 QN 182  
DB 181 QN 182

## RESULT 3

ID AAW37448 standard; protein; 200 AA.

AC AAW37448;

DT 08-JUN-1998 (first entry)

DE Tick vasocactive amine binding protein 1 MS-HBPI.

XX Male-specific vasocactive amine binding protein 1; MS-HBPI;

KW histamine; serotonin; assay; antihistamine; anti-inflammatory;

KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;

XX transgenic animal; tick.

XX Rhipicephalus appendiculatus.

OS Key Location/Qualifiers

FT Peptide 1..18

FT Modified-site 79..81

FT /note="Asn is N-glycosylated"

FT WO9744451-A2.

PN 27-NOV-1997.

PD 19-MAY-1997; 97WO-GB01372.

PE 18-APR-1997; 97GB-0007844.

PR 18-MAY-1996; 96GB-0010484.

PA (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC;

XX WPI; 1998-018506/02.

DR N-PSDB; AAV00229.

XX New vasocactive amine binding proteins and related nucleic acid,

PT vectors - transformed cells and transgenic animals, used for

PT assaying or removing histamine and as antihistamine or

PT anti-inflammatory agents

XX Example 2; Fig 3; 44pp; English.

PS This protein comprises tick Rhipicephalus appendiculatus (Ra) novel

XX male-specific histamine binding protein 1 (MS-HBPI). Its amino

CC acid sequence was deduced from a cDNA clone (see AAV00227) obtained

CC from a salivary gland cDNA library. Female-specific HSP1 and



PT conjunctivitis, or for manufacturing a medicament for treating or  
 PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
 PT conjunctivitis  
 PS Claim 4; Page 5-6; 19pp; English.  
 XX  
 CC The present invention relates to the use of a histacalin protein  
 CC for treating or preventing conjunctivitis. The present sequence  
 CC is the histacalin protein FS-HBP1. The invention is particularly useful  
 CC in the treatment of allergic or seasonal conjunctivitis.  
 CC  
 SQ Sequence 171 AA;  
 Query Match 33.5%; Score 331.5; DB 22; Length 171;  
 Best Local Similarity 42.0%; Pred. No. 3.6e-26;  
 Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
 QY 2 PTANENKLGSYQAMKSLQODQKRYIAQATOTGTGWCGETCVSYTAERKIG--KKK 59  
 Db 3 PDMADEANGAHQDAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDEKS 62  
 QY 60 LNAITLYKNKHLTDLKSHETITVWKAYDYTTENGIRYETQGTOTFEDEVFVPSDYKNC 119  
 Db 63 IQAEFLMNNADTMQKATEKVTAKYKGYRENAFRETED--GQVFTDVIAYSD--DNC 119  
 QY 120 DVIFVPERGSDGDEYELMWSEDKIDKIPDCKFTMAVFAQQQKFTVRYNNTDSSC 175  
 Db 120 DVIVPGTGNNEG-YELMTTD--YDNI PANCLNKFNRYAVGRE--TRDVFET--SAC 169  
 RESULT 6  
 AAB73261  
 ID AAB73261 standard; protein; 171 AA.  
 AC AAB73261;  
 XX  
 DT 15-MAY-2001 (first entry)  
 XX  
 DE Histacalin protein FS-HBP2.  
 XX  
 KW FS-HBP2; histacalin protein; antiinflammatory; antiallergic;  
 KW ophthalmological; allergic rhinitis.  
 XX  
 OS Unidentified.  
 OS  
 PN WO200116164-A2.  
 PN  
 PD 08-MAR-2001.  
 PD  
 PF 24-AUG-2000; 2000WO-GB03287.  
 PF  
 PR 01-SEP-1999; 99GB-0020673.  
 PR  
 PA (EVOL-) EVOLUTEC LTD.  
 PA  
 PI Nuttall PA, Paesen GC;  
 PI  
 DR WPI; 2001-218521/22.  
 DR  
 XX  
 PT Use of histacalin proteins for treating or preventing allergic  
 PT rhinitis, or for manufacturing a medicament for treating or preventing  
 PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis  
 XX  
 XX Disclosure; Pages 4-6; 19pp; English.  
 XX  
 CC The present invention relates to a method for treating or preventing  
 CC allergic rhinitis. The method involves employing a blood-feeding  
 CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
 CC sequence is one such histacalin protein. The histacalin protein, is  
 CC useful for treating or preventing allergic rhinitis, both seasonal and  
 CC perennial allergic conjunctivitis.  
 CC  
 SQ Sequence 171 AA;

Query Match 33.5%; Score 331.5; DB 22; Length 171;  
 Best Local Similarity 42.0%; Pred. No. 3.6e-26;  
 Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
 QY 2 PTANENKLGSYQAMKSLQODQKRYIAQATOTGTGWCGETCVSYTAERKIG--KKK 59  
 Db 3 PDMADEANGAHQDAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDEKS 62  
 QY 60 LNAITLYKNKHLTDLKSHETITVWKAYDYTTENGIRYETQGTOTFEDEVFVPSDYKNC 119  
 Db 63 IQAEFLMNNADTMQKATEKVTAKYKGYRENAFRETED--GQVFTDVIAYSD--DNC 119  
 QY 120 DVIFVPERGSDGDEYELMWSEDKIDKIPDCKFTMAVFAQQQKFTVRYNNTDSSC 175  
 Db 120 DVIVPGTGNNEG-YELMTTD--YDNI PANCLNKFNRYAVGRE--TRDVFET--SAC 169  
 RESULT 7  
 AAM37447  
 ID AAM37447 standard; Protein; 190 AA.  
 AC AAM37447;  
 XX  
 DT 08-JUN-1998 (first entry)  
 XX  
 DE Tick vasoactive amine binding protein 2 FS-HBP2.  
 XX  
 KW Female-specific vasoactive amine binding protein 1; FS-HCP1;  
 KW histamine; serotonin; assay; antihistamine; anti-inflammatory;  
 KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
 KW transgenic animal; tick.  
 XX  
 OS Rhipicephalus appendiculatus.  
 OS  
 FH  
 FX Key Location/Qualifiers  
 FT Peptide 1..19  
 FT /label= Sig-peptide  
 XX  
 PN WO9744451-A2.  
 PN  
 PD 27-NOV-1997.  
 PD  
 PF 19-MAY-1997; 97WO-GB01372.  
 PF  
 PR 18-APR-1997; 97GB-0007844.  
 PR  
 PR 18-MAY-1996; 96GB-0010484.  
 PR  
 PA (OXFO-) OXFORD VACS LTD.  
 PA  
 PI Nuttall PA, Paesen GC;  
 PI  
 DR WPI; 1998-018506/02.  
 DR  
 DR N-PSDB; AAV00228.  
 DR  
 XX  
 PT New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 XX  
 PS Example 2; Fig 2; 44pp; English.  
 PS  
 XX  
 CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
 CC female-specific histamine binding protein 2 (FS-HBP2). Its amino  
 CC acid sequence was deduced from a cDNA clone (see AAV00228) obtained  
 CC from a salivary gland cDNA library. FS-HBP1 (see AAM37446) and  
 CC male-specific HSP1 (see AAM37448) and a related protein, D.KRT6 (see  
 CC AAM37449) from Dermacentor reticulatus, were also identified. These  
 CC novel vasoactive amine binding proteins (VABPs) can be expressed  
 CC in host cells using e.g. a baculovirus expression system. They can  
 CC be used: (i) to assay histamine (or other VA such as serotonin) in  
 CC body fluids or cell culture supernatants, e.g. to monitor the  
 CC effect of allergens; (ii) for binding VA, e.g. to remove histamine







Query Match	27.0%	Score 267.5;	DB 20;	Length 190;
Best Local Similarity	39.3%	Pred. NO. 1.8e-19;		
Matches 70: Conservative	20;	Mismatches 77;	Indels 11;	Gaps 7.

CC	vacuoles directed against blood sucking ectoparasites.
XX	
SQ	Sequence 203 AA;

XX	Sequence	203 AA:	26.6%;	Score 263;	DB 20;	Length 203;
SC	Query Match		35.1%;	Pred. No. 5,8e-19;		
XX	Best Local Similarity		27;	Mismatches 57;	Indels 36;	Gaps 9
XX	Matches	65;	Conservative			
QY	1 NPTWANEKLCISYDAMKSLDODNKKRYLLAQAOTTTGWCAGEETCYSYVA----	EKIG 56				
DB	32 NTTW-HSKELNYYODAMKSNONSTYTYFFLRSTFNNDNSVYMGKNTCLSVTYSKHESTF	90				
QY	57 KKKLNATLLYKN-----KHLIDKESHETITVWKRAYDYTTENGIRKEYEQGRITQTF	107				
DB	91 TVEYITTT--YKNOSQOWSMTEBNATVAOE--EGYVKNKINIOFTTENTNRK-----F	136				
QY	108 EDVEFSDYKKKNDVIFVP-KERGSDEGYIELVWSBDKIDKIPDCKFTMAVFAQOQERTV	166				
DB	137 NDTVFTFGQYCCDDLIYRYKENG-----YELWVRSDYLIQNPPTCCQFIPLDLVA--LGRTT	189				
QY	167 RNVYT 171					
DB	190 YNIST 194					
XX	RESULT 15					
XX	AAM37449					
XX	ID AAM37449 standard; Protein; 209 AA.					
XX	AC AAM37449;					
XX	AA 08-JUN-1998 (first entry)					
XX	DE Tick vasoactive amine binding protein D. RET6.					
XX	Vasoactive amine binding protein; D. RET6; histamine; serotonin;					
XX	KM assay; antihistamine; anti-inflammatory; insect bite; snake bite;					
XX	KM scorpion bite; dermatitis; vaccine; transgenic animal; tick.					
XX	OS Dermacenter reticularis.					
XX	Key location/Qualifiers					
XX	FM 1.28					
XX	FT Peptide /label- Sig-peptide					
XX	PN WO9744451-A2.					
XX	PD 27-NOV-1997.					
XX	PF 19-MAY-1997; 97WO-GB01372.					
XX	PR 18-APR-1997; 97GB-0007844.					
XX	PR 18-MAY-1996; 96GB-0010484.					
XX	PA (OXFO-) OXFORD VACS LTD.					
XX	PI Nuttall PA, Paesen GC;					
XX	PI MPI: 1998-018506/02.					
XX	DR N-PSDB; AAV00230.					
XX	PT New vasoactive amine binding proteins and related nucleic acid,					
XX	PT vectors - transformed cells and transgenic animals, used for					
XX	PT assaying or removing histamine and as antihistamine or					
XX	PT anti-inflammatory agents					
XX	Example 2; Fig 4; 44pp; English.					
XX	This protein comprises tick Dermacenter reticularis (Dr) novel					
XX	CC vasoactive amine binding protein (VABP) D. RET6. Its amino acid					
XX	CC sequence was deduced from a cDNA clone (see AAV00230) obtained					
XX	CC from a Dr salivary gland cDNA library. 3 Novel VASPs, designated					
XX	CC FS-HBP1, FS-HBP2 and MS-HBP1 (see AAM37446-48), of the tick					
XX	CC Rhipicephalus appendiculatus have also been identified. The					

VABPs can be expressed in host cells using e.g. a baculovirus expression system. They can be used: (i) to assay histamine (or other VA such as serotonin) in body fluids or cell culture supernatants, e.g. to monitor the effect of allergens; (ii) for binding VA, e.g. to remove histamine from blood, food, cell cultures etc.; (iii) as an antihistamine or anti-inflammatory agents, e.g. for treating insect, snake or scorpion bites or dermatitis, or as a carrier for slow release of histamine-related compounds; (iv) in vaccines to protect against metazoan parasites especially in animals; (v) as reagents for studying inflammation, involvement of VA in ulcer formation or the immune response etc. VABPs provide a more sensitive assay for histamine than low-affinity antibodies currently used. They may also be more effective and safer than conventional antihistamines.

Query Match	25.4%	Score 251.5;	DB 19;	Length 209;
Best Local Similarity	35.3%	Pred. No. 9, 3e-18;		
Matches 61; Conservative	29;	Mismatches 66;	Indels 17;	Gaps 8;

```

OY 1 NPNMAEAKGSGVODAMKSLDODKKRYLKOATOTOT-GWGSEFTCVSTAKIRKK 59
Db 30 NPNMAHEELLKGQODAMKSIDGVSATVLAATTENTGSMGSOFCLOY--QEIIRKE 87
OY 60 LNATI---LYKKHLTDLKESHETITWAKVDY---TTENGKIVENOGTPTOFFEDV 112
Db 88 EDTTVSVTFERNAS-SPKIYNTYETKANPQYKIRIRNAIEQVGGGLNTI--DTLI 144
OY 113 FSDYKACVDIENPKEKSGDEGDYELWASEDKTIDKTPDCCKTPTMAFPAQOQK 165
Db 145 FTQELGELDFFVVP---NADQG--CELMVKKSHYKHHPDCTVFNENFCAKDKKT 193

```

Search completed: August 1, 2003, 12:22:49  
Job time : 57.5067 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 49.1829 seconds  
(without alignments)  
897.203 Million cell updates/sec

Title: US-10-087-195-7

Perfect score: 936  
Sequence: 1 NQPDMADEANGAHQDAWKS.....FNEYAVGRETRDVFYSACLE 171

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	282	30.1	209	5	Q8WSK7
2	114.5	12.2	311	5	Q8WVC2
3	104.5	11.2	1898	5	Q8WVC2
4	103	11.0	1590	2	Q59983
5	97	10.4	1590	2	Q59983
6	96.5	10.3	1449	5	Q59983
7	92.5	9.9	944	3	Q8WVC2
8	92	9.8	176	5	Q8WVC2
9	91.5	9.8	221	5	Q8WVC2
10	91.5	9.8	386	16	Q8WVC2
11	91.5	9.8	640	8	Q8WVC2
12	91	9.7	306	5	Q8WVC2
13	91	9.7	450	16	Q8WVC2
14	90.5	9.7	1083	5	Q8WVC2
15	89.5	9.6	542	3	Q8WVC2
16	89.5	9.6	845	10	Q8WVC2

17	88	9.4	606	16	Q977Y3
18	87	9.3	377	12	Q90B34
19	87	9.3	193	5	Q8IDK4
20	86.5	9.2	193	12	Q8IBS3
21	86.5	9.2	254	5	Q23871
22	86.5	9.2	723	16	Q8P1H6
23	86.5	9.2	736	16	Q9A0E7
24	86.5	9.2	736	16	Q8K7Z8
25	86	9.2	6077	5	Q8IC56
26	85.5	9.1	327	16	Q8CP44
27	85.5	9.1	573	5	Q8IDM2
28	85.5	9.1	617	16	Q8X1I6
29	85.5	9.1	686	8	Q8WVG0
30	85.5	9.1	1237	16	Q92CC8
31	85.5	9.1	1612	16	Q8Y591
32	84	9.0	2488	5	Q813K5
33	83.5	8.9	4494	5	Q815I2
34	83	8.9	601	2	Q45821
35	83	8.9	874	2	Q9S3L0
36	82.5	8.8	220	5	Q8WVB8
37	82.5	8.8	496	5	Q9CSW0
38	82	8.8	419	10	Q9FJH4
39	81.5	8.7	374	16	Q8CX67
40	81.5	8.7	897	5	Q8ID14
41	81.5	8.7	1026	2	Q8L2H9
42	81.5	8.7	1129	5	Q9YEP3
43	81.5	8.7	1956	5	Q81IE1
44	81.5	8.7	3597	5	Q81LR5
45	81	8.7	1263	11	Q9QY22

## ALIGNMENTS

RESULT 1  
Q8WSK7 PRELIMINARY: PRT; 209 AA.  
ID Q8WSK7  
AC Q8WSK7  
DT 01-MAR-2002 (TRENBLREL. 20, Created)  
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
DE Serotonin and histamine binding protein.  
OS Dermacentor reticulatus.  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Ixodidae; Dermacentor.  
ON NCBI\_TaxID=57047;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sangamadech S., Paesen G.C., Nuttall P.A.;  
RT "A high affinity serotonin- and histamine-binding 11pocalin secreted by blood-feeding ticks.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF217101; AAL56644.1;  
DR InterPro: IPR002970; His\_binding.  
DR Pfam: PF02098; His\_binding; 1.  
SQ SEQUENCE 209 AA; 23889 MW; E5EB8BE710859591 CRC64;

Query Match 30.1%; Score 282; DB 5; Length 209;  
Best local similarity 37.5%; Pred. No. 5.2e-17;  
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;  
Db 3 PMADEANGAHQDAWKS...ATEKYAVKMGY-NRENAFRETEDGQFTVIVASD-D 117  
62 SIQAEFLPMNNDITNQF--ATEKYAVKMGY-NRENAFRETEDGQFTVIVASD-D 117  
91 TTSVFTF-RNASSPIKRYNTEYKAVQYGYKRNRIEYQVGGLITDTLFTTDE 149  
QY 118 NCDVITYVGTDSNEGCELYMT--TDYDNPANCLKNFXY-AVGRETRDVFYSACL 170  
DB 150 LCDVYVPMAD---GCELMVKKSHKHYDCTFVFNVCCKRDRTYDIFNEECV 202

```

RESULT 2
ID 08MVC2 PRELIMINARY; PRT; 311 AA.
AC 08MVC2;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DE Putative secreted histamine binding protein.
DE Ixodes scapularis (Black-legged tick) (Deer tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodidae; Ixodidae; Ixodes.
OX NCBI_Taxid=6945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rhode Island; Tissue-Salivary gland;
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,
RA Mather T.N., Ribeiro J.M.C.;
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes
RT scapularis."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF483718; AAM93640.1;
SQ SEQUENCE 311 AA; 36581 MW; D7E529FEFE4CFA9A CRC64;

Query Match 12.2%; Score 114.5; DB 5; Length 311;
Best Local Similarity 22.8%; Pred. No. 0.041;
Matches 42; Conservative 26; Mismatches 63; Indels 53; Gaps 7;

QY 1 NOPDMADEAANGAHQDAMKSLKADVENYVVKATYKNDPYWG-----NDFTCGVAND 55
DB 31 NNPD-LNRKDLGAMQDAMRIKPTANHSYLIYS-----GNGTREHYEDVRCLOVHSSD 84
QY 56 VNDEKSIQAEFLFNADTNMOPATEKVTAVKMYGYNRENA----- 97
DB 85 LNTYTKSANTSKYNTSKRMNSTQYOAOKQYSIENIMHIGPOREVTSPNGTCY 144
QY 98 ---FRYTEEDG-----QVFTDVIAYSNDNDVITYPTDNGEEY---EL 136
DB 145 NLNENFLCESGGCRTHHCEWQKRWYTSERYLESTPLC---YVVASLDDDEGESCEF 201
QY 137 WTTD 140
DB 202 WLSE 205

RESULT 3
ID 08IL35 PRELIMINARY; PRT; 1898 AA.
AC 08IL35;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Hypothetical protein, conserved.
DE Pf14_0414.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-3D7;
MDLINE=22255705; Pubmed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.D., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Garucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).

```

```

DR EMBL; AE014823; AAN37027.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1898 AA; 223829 MW; 7B5CB02A7091822F CRC64;

Query Match 11.2%; Score 104.5; DB 5; Length 1898;
Best Local Similarity 27.5%; Pred. No. 3;
Matches 44; Conservative 23; Mismatches 56; Indels 37; Gaps 9;

QY 17 AMKSLKADVENYVVKATYKN-----DPVGNDFTCGVANDVNEDEKSIQAEFLF 69
DB 571 AMKSSVLEB---YCKSKYENKKDLIDYNSLFT--HYEN--NHDSKEQSNFF 623
QY 70 MNNA-----DTNQFATEKYAVKMYGYN--ENAFRYTEQVFTDVI-----AY 114
DB 624 HNNANSSYTFEDTISENF-----ISKYFPRISYKKSYPDFEKEKFDNLITGCMCY 676
QY 115 SDQNCVDIYVPGTDGNEGEELWTTVDNIPANLKNFNE 154
DB 677 IYDN-KIVLNDYKGIOTLFKKRYDYDNIIFSVCINNRQ 715

```

```

RESULT 4
ID 059983 PRELIMINARY; PRT; 1590 AA.
AC 059983;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE Glucosyltransferase-I precursor (Ec 2.4.1.5).
DE GTFI.
GN Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OM2176;
MDLINE=94146405; Pubmed=8312602;
RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d
RT Streptococcus sobrinus."
RL DNA Seq. 4:19-27(1993).
DR EMBL; D13858; BAA02876.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glucosyltransferase; Signal; Transferase.
FT SIGNAL 1
FT CHAIN 39 1590
FT SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;

Query Match 11.0%; Score 103; DB 2; Length 1590;
Best Local Similarity 25.9%; Pred. No. 3.2;
Matches 52; Conservative 19; Mismatches 54; Indels 76; Gaps 12;

QY 2 QPDM-----ADEAANGAHQ--DAWKSILKADVENYVYV-----KATYK-N 38
DB 335 QPQNGSEKRPYDDHLDNGALFDNCTDLPDQSNRYLRPTNQTGSLDSFFTYNPN 394
QY 39 DPVWGNDFTCGVANDVNEDEKSIQAEFL-----FMNADTNMOPATEKVT 86
DB 395 DPLGGVDE---LLANDVNSNPVQAEQLMWLHYLLNFGSITANDADAN--FDSIRVDA 448
QY 87 VKMYGYRENAFRFTEEDGVFTDVIAYSNDNDVITYPTDNGEEG-----YELMT- 138
DB 449 V-----DAMDADLDIOISSDYLKAAV--GIDNNKNNANHHVSIIVEAMSD 489
QY 139 -----TDYDNIPANLKNF 152
DB 490 NDTPYLLHDDGDNL-MNMDNKF 509

```

RESULT 5  
ID 055263 PRELIMINARY; PRT; 1590 AA.  
AC 055263;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE GTF-I.  
GN GLUCOSYLTRANSFERASE.  
OS Streptococcus sobrinus.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1310;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 33478;  
RA Sato S.;  
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase  
RT produced from Streptococcus sobrinus ATCC 33478."  
RL Am. Kagoshima Univ. Dental School 16:23-29(1996).  
DR EMBL: D63570; BAA09792.1; -  
DR InterPro: IPR002479; CW\_binding.  
DR InterPro: IPR003318; Glyco\_hydro\_70.  
DR Pfam: PF01473; CW\_binding\_1; 15.  
DR Pfam: PF02324; Glyco\_hydro\_70; 1.  
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match 10.48; Score 97; DB 2; Length 1590;  
Best Local Similarity 24.98; Pred. No. 11;  
Matches 50; Conservative 22; Mismatches 54; Indels 76; Gaps 12;

QY 2 QPQW-----ADEANGAHQ-DAMSLKADVENYVYV-----KATYK-N 38  
DB 335 QPQWNGSEKRYDHLQNGALKFPNOSDLPTDQSNRLNKRPTNGSLDSFTYNNAN 394  
QY 39 DPVWNGDFTCGVYVANDVNEDEKSIQAEFL-----FNNDPTNNQFATEKYTA 86  
DB 395 DPLQGYEF-----LLANDVNSNPVQAQEQALMHLHYLLNFGSIVAKDADAN--FDSIRYDA 448  
QY 87 VKMVGYNRENAFRETEGQFTYVIAVSDNCVYIYVPGDGAEEG-----YELMT- 138  
DB 449 V-----DNDVADLLQISSDYLKAAV--GIDKNNKANNHVSIVEAMSD 489  
QY 139 -----TDYDNIPANCLNKF 152  
DB 490 NDTPYLHDDGDNL-MNMDNKF 509

RESULT 6  
ID 09GSZ7 PRELIMINARY; PRT; 1449 AA.  
AC 09GSZ7;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Carbamoyl phosphate synthetase II (EC 1.4.3.6) (Copper amine oxidase)  
DE (Fragment).  
OS Plasmodium berghei.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K173;  
RA Karib M.T., Davies N.P., Stewart T.S.;  
RT "Unique insert Sequences within Carbamoyl Phosphate Synthetase II Gene  
RT from Plasmodium berghei and Plasmodium chabaudi."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +  
CC H(2)O(2).  
CC -1- COPACITOR: BINDS 1 COPPER ION AND 1 TOPAQUINONE PER SUBUNIT (BY  
CC SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE COPPER/TOPAQUINONE OXIDASE FAMILY.  
DR EMBL: AF286897; AAC10785.1; -

DR HSSP; P00968; ICS0.  
DR InterPro: IPR005483; CPase\_L.  
DR InterPro: IPR005479; CPase\_L\_D2.  
DR InterPro: IPR005480; CPase\_L\_D3.  
DR InterPro: IPR005481; CPase\_L\_N.  
DR InterPro: IPR001317; CPS\_GATase.  
DR InterPro: IPR000269; CUNH\_oxidase.  
DR InterPro: IPR000991; GATase.1.  
DR Pfam: PF00289; CPase\_L\_chain; 1.  
DR Pfam: PF02786; CPase\_L\_D2; 1.  
DR Pfam: PF02787; CPase\_L\_D3; 1.  
DR Pfam: PF00117; GATase.1.  
DR PRINTS: PR00098; CPase.1.  
DR PRINTS: PR00099; CPase.1.  
DR PRINTS: PR00096; GATase.  
DR PROSITE: PS01164; COPPER\_AMINE\_OXID\_1; 1.  
DR PROSITE: PS00867; CPase\_2; 1.  
KW Copper; Oxidoreductase; TPO.  
FT NON\_TER 1  
FT TER 1  
SQ SEQUENCE 1449 AA; 167415 MW; CC72C27C72077F6B CRC64;

Query Match 10.38; Score 96.5; DB 5; Length 1449;  
Best Local Similarity 19.68; Pred. No. 11;  
Matches 37; Conservative 36; Mismatches 65; Indels 51; Gaps 9;

QY 6 ADEANGAHQDAMSLKAD-----VENVYVYKATYKNDPVWNGDFTCGVYVANDV 56  
DB 1202 SEDEENVNPD--KSMRRDDNRGINNIANNSTYI-----RDSVYNNEXK-IKKMRELI 1252  
QY 57 NEDEKS--IQAEFLFNNDPTNNQFATEKYVAVMYGVNRENAFRETEGQVPTD---- 110  
DB 1253 NNDKSEVIAKSKFANSTNSKCEYINSNKKAKIYNNDNTVYNTNDEEENKNSLS 1312  
QY 111 -----VIAVSDN-----CDVIYVPTDGN-----EGCYELMTTQYNNIPANC 148  
DB 1313 NNNEMKNNNNLINSYNDNRKSDCFSELSTYQNHKKRINDIENDY-----GYEDLYSGS 1367  
QY 149 LKRFNEYAV 157  
DB 1368 DNTSYSTSI 1376

RESULT 7  
ID 09HE41 PRELIMINARY; PRT; 944 AA.  
AC 09HE41;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Related to importin beta-2 subunit (Transportin).  
GN B9B15.030.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hohnselt J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL451014; CAC18173.2; -  
DR HSSP; Q92973; IOBK.  
DR InterPro: IPR001494; Importinb\_N.  
DR PROSITE: PS50166; IMPORTIN\_B\_NT; 1.  
SQ SEQUENCE 944 AA; 104908 MW; B5C48DCAE470D5EF CRC64;

Query Match 9.98; Score 92.5; DB 3; Length 944;  
Best Local Similarity 23.28; Pred. No. 14;





DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE MDH dehydrogenase subunit F (Fragment).  
GN MDH.  
OS Lyimachia nemorum (yellow pimpernel).  
OC Chloroplast.  
OC Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; Ericales; Myrsinaceae; Lysimachia.  
OX NCBI\_TaxID=110756;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kallersjö M., Bergqvist G., Anderberg A.A.;  
RT "Genetic reassignment in primulaid families of the Ericales s. l.  
(Angiosperms): a phylogenetic analysis based on DNA sequences from  
three chloroplast genes and morphology.";  
RL Am. J. Bot. 0:0-0(2000).  
DR EMBL; AF213747; AAC43878.1;  
DR InterPro: IPR003916; NADHox\_oxrd5.  
DR InterPro: IPR002128; Oxidored\_q1.  
DR InterPro: IPR001516; Oxidored\_q1\_C.  
DR Pfam: PF00361; Oxidored\_q1\_1.  
DR Pfam: PF01010; Oxidored\_q1\_C\_1.  
DR Pfam: PF00662; Oxidored\_q1\_N\_1.  
DR PRINTS: PR01434; NADHoxGNASE5.  
KM NAD, Oxidoreductase; Plastocyanine; Chloroplast.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 640 AA; 72317 MW; 120FA9DAE50E09E5 CRC64;

Query Match 9.8%; Score 91.5; DB 8; Length 640;  
Best Local Similarity 21.9%; Pred. No. 10;

Matches 33; Conservative 25; Mismatches 58; Indels 35; Gaps 5;

QY 27 NVY---YVVKATYKNDPFGNDFTCVGMANDVNEDEKSIQAEFLFMNN----- 72  
DB 421 NVYFKNNNGKTSHTSISLWCK-----GNSKKKKONIFLLNNNESTLKKRI 468  
QY 73 --ADTNKQFATEKVAVKMYGYNRENAFRYETEDGQVFTDIYASDNCQVYIYPGTGN 130  
DB 469 YRIDENKQKMTFRPFTITSTFGNKTYSYPESDNTMLFTLLIFLFTSC--VGFIGIPFN 526  
QY 131 EEEVEL-----WTTDVNIPANCLNKNEYA 156  
DB 527 QEGTDLDTLSKWLTPSTINLHONLKNYDWS 557

## RESULT 12

ID 08MVC3 PRELIMINARY; PRT; 306 AA.  
AC 08MVC3;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OS Eukaryote: Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
RA Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis.";  
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF483717; AAM93699.1;  
SQ SEQUENCE 306 AA; 34232 MW; B2DB0910CC4395D CRC64;

Query Match 9.7%; Score 91; DB 5; Length 306;

Best Local Similarity 24.1%; Pred. No. 4.6;  
Matches 35; Conservative 21; Mismatches 65; Indels 24; Gaps 4;

QY 7 DEANGAHQDAWMSLKADVENYVVKATYKNDPFGNDFTCVGMANDVNEDEKSIQAE 66  
DB 34 NEPSLGPLOASWAKINSSKDDPFLMFRSRNHEP-----NITCVVVTSLNETLKIYFT 89  
QY 67 FLFNNNADTNMVFATEKVAVKMYGYNRENAFR-----YETEDGQVFTDIYASDD 117  
DB 90 RTYINETDGNLDLEVOYRALNQTLYKLENYIRAGLGTSPDKPTPLGS--NMITYGDY 147  
QY 118 NCDVIYVPGTD-----GNEEG 133  
DB 148 SCNTSSKPLDMDLKAADAVGSEAG 172

## RESULT 13

ID 09PQ49 PRELIMINARY; PRT; 450 AA.  
AC 09PQ49;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein U0441.  
GN U0441.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Serovar 3;  
RX MEDLINE-20500219; PubMed-11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Helner C.R., Chen E.Y.,  
RA Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum.";  
RL Nature 407:757-762(2000).  
DR EMBL; AEO02140; AAF30853.1;  
DR InterPro: IPR000209; Peptidase\_S8.  
DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 450 AA; 52947 MW; BD0BAB5D1FBA6206 CRC64;

Query Match 9.7%; Score 91; DB 16; Length 450;  
Best Local Similarity 23.4%; Pred. No. 7.5;  
Matches 37; Conservative 23; Mismatches 60; Indels 38; Gaps 8;

QY 7 DEANGAHQDAWMSLKADVENYVVKATYKNDPFGNDFTCVGMANDVNEDEKS 62  
DB 19 DEISENSHIKIKWKLMDLQNDVKIINHSGWSP---NDFLNIGFYNPINKKEIN-DDKD 74  
QY 63 IQAEFL-----FMNNAADTNMVFATEKVAVKMY-----GYRENAFRYE 101  
DB 75 FKRELKRYIDTLNKDDIVNNTINKKFNKYIDFISYFDIINIFSGNSYND--IRYR 132  
QY 102 TEGDQV---FTDIYAS---DONCDVIYVPGTGNEE 132  
DB 133 NKNGHLEINWFDISKYSFDRFNQNSIYVGSTKNNE 170

RESULT 14

ID 08I3G1 PRELIMINARY; PRT; 1083 AA.  
AC 08I3G1;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN PF0133W.  
OS Plasmodium falciparum (Isolate 3D7).  
OS Eukaryote: Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,  
RA Hall N., Bowman S., Churcher C., Quail M., Barrell B.,  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE-2225706; PubMed-12368867;

RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,

RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,

RA Bucke C.O., Burrows C., Cherevach I., Chillingworth C.,

RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C.,

RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,

RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,

RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,

RA Humphray S., Jagels K., James K.D., Johnson D., Kermorou A.,

RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,

RA Line A., Maddison M., McLean J., Mooney P., Moutle S., Murphy L.,

RA Oliver K., Ormond D., Price C., Quail M.A., Rabinovitch E.,

RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,

RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,

RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,

RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;

RT "Sequence of plasmidium falciparum chromosomes 1, 3-9 and 13.",

RT Nature 419:527-531(2002).

DR EMBL: AL929354; CAD51668.1; -.

KW Hypothetical protein.

SV SEQUENCE 1083 AA; 128123 MW; 487F49816957C623 CRC64;

Query Match 9.7%; Score 90.5; DB 5; Length 1083;

Best Local Similarity 23.5%; Pred. No. 25;

Matches 36; Conservative 25; Mismatches 49; Indels 43; Gaps 7;

QY 25 VENVYVYKATYKNDPVWGNDFTC-----VGYMANDVNEDEKSIQAEFL-FMN-----N 72  
Db 506 INNMFEIKIHNSNDKTKKFCSEYIIISINILINKIIECNTTYENYILEFLDLFEKN 565  
QY 73 ADNNMQFATEKYTA-VKMTYENRENAFR-----YETEDGVFTDVIAYSDDNCVY 122  
Db 566 ACVYSKDNTKEKYIKYKKKKRKNCFENSYKWDIEFYNIED-----LYINDSKVYI 617  
QY 123 YVPGTDGNEEGYELMTTDDYDINPANCINKNEY 155  
Db 618 YINGKEKE-----GANRNTYIDKY 637

RESULT 15

Q9P378

ID 09P378

AC 09P378;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Protein similar to yeast gas1 glycopospholipid-anchored surface

DE glycoprotein, putative glycosidase that regulates the crosslinking of

DE beta-1,6-glucans in the cell wall.

GN SPAC19B12.02C.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972h-;

RA Wood V., Rajandream M.A., Barrell B.G., Seeger K., Harris D.,

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL390814; CAC000550.1; -.

DR GeneDB\_Spombes; SPAC19B12.02c; -.

DR InterPro; IPR004886; GAS1.

DR Pfam; PF03198; GAS1; 1.

SV SEQUENCE 542 AA; 58115 MW; 1D243FDA5AD2EC3A CRC64;

Query Match 9.6%; Score 89.5; DB 3; Length 542;

Best Local Similarity 25.1%; Pred. No. 13;

Matches 44; Conservative 22; Mismatches 60; Indels 49; Gaps 10;

QY 13 AHQDAKSLKADYENYVYKATYKNDPVWGND-----TCVGMANDVN 57  
Db 102 AFQDAGIYVLSLAQPY---EALSSDPTVTVDLFSRYTEVDLSLAPYDNMLFFIAG--N 156  
QY 58 EDEKSIQAEFLFMNADTN-MQFATEKVTAVKMY---GYNRENAFRETEDGVFTDVY 112  
Db 157 E-----VIQNTNTNAAFYKAAVRDKSYIKSSGY-RQIPVGYSTNDEEYTRDPM 206  
QY 113 AY-----SDNCDVI-----YVPGTDGNEEGYELMTTDDYDINPANCINKNEY 155  
Db 207 AYVFDCGDDDDHVDYFGINITYEWCQSDSFVSSGYQERTTEFSNMTVPMI--FSEF 259

Search completed: August 1, 2003, 12:19:59  
Job time : 51.1829 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: August 1, 2003, 12:15:03 ; Search time 10.4229 Seconds

(without alignments)  
771.532 Million cell updates/sec

Title: US-10-087-195-7  
Perfect score: 936  
Sequence: 1 NOPMDAENGANGHQAQDAKMS.....FNEYAVGRTFVTSACLE 171

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	100.0	190	1	HBP2_RHIAP
2	630	67.3	190	1	HBP1_RHIAP
3	331.5	35.4	200	1	HBP2_RHIAP
4	105	11.2	1592	1	GTF2_STRDO
5	91	9.7	1597	1	GTF1_STRDO
6	89.5	9.6	542	1	YLM2_SCHPO
7	84.5	9.0	858	1	CYAG_DICDI
8	84	9.0	624	1	AMYG_AKXAD
9	83.5	8.9	404	1	CT99_HUMAN
10	83.5	8.8	2911	1	FN22_HUMAN
11	83.5	8.8	954	1	XYNA_RUMFL
12	81.5	8.7	1129	1	RPA2_DROME
13	80	8.5	1695	1	CICA_PAEPP
14	79.5	8.5	2907	1	FN22_MOUSE
15	78.5	8.4	1822	1	ITB4_HUMAN
16	77.5	8.3	421	1	DHE2_PEPAS
17	77.5	8.3	512	1	HEX_ADE09
18	77.5	8.3	1026	1	BGAL_STRTR
19	77	8.2	335	1	Y687_METTA
20	77	8.2	365	1	VG13_BP803
21	76	8.1	393	1	VPAP_HSV6U
22	76	8.1	393	1	VPAP_HSV6U
23	75.5	8.1	489	1	INGR_HUMAN
24	75.5	8.1	1279	1	APU_THESA
25	75	8.0	265	1	YRRA_MYCCA
26	75	8.0	557	1	PGMU_EMEHI
27	74	7.9	410	1	CD66_HALN1
28	73.5	7.9	517	1	GUWA_CLOLO
29	73.5	7.9	541	1	VA97_MYCPN
30	73.5	7.9	828	1	PMFC_PROMI
31	73	7.8	625	1	TAP2_HAEIN
32	73	7.8	749	1	SPOT_SPICI
33	73	7.8	960	1	DIG1_DROME

34	72.5	7.7	282	1	POR2_XENLA
35	72.5	7.7	661	1	PSO2_YEAST
36	72.5	7.7	846	1	Y661_MYCTU
37	72.5	7.7	889	1	IRE1_RAT
38	72	7.7	283	1	PHUD_BACE
39	72	7.7	363	1	OMPE_SALTI
40	72	7.7	600	1	ABRA_PLAVG
41	72	7.7	603	1	BGLR_ECOLI
42	72	7.7	743	1	ABRA_PLARC
43	72	7.7	1041	1	UN83_CAEEL
44	72	7.7	1227	1	CIBE_BACTU
45	71.5	7.6	506	1	CYSP_PLAVN

## ALIGNMENTS

RESULT 1					
ID	HBP2_RHIAP	STANDARD:	PRT:	190 AA.	
AC	077421:				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Female-specific histamine-binding protein 2 precursor (FS-HBP2).				
OS	Rhipicephalus appendiculatus (Brown ear tick).				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OC	Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.				
OX	NCBI_TaxID=34631;				
RM	(1)				
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).				
RC	TISSUE-Salivary gland:				
RX	MEDLINE=99286454; PubMed=10360182;				
RA	Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;				
RT	"Tick histamine-binding proteins: isolation, cloning, and three-				
RL	dimensional structure".				
CC	Mol. Cell 3:661-671(1999).				
CC	-1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO				
CC	OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO				
CC	SUPPRESS INFLAMMATION DURING BLOOD FEEDING.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL: U96081; AAC63107.1; -				
DR	PDB: 1OFT; 19-APR-00.				
DR	PDB: 1OFV; 19-APR-00.				
DR	InterPro: IPR002970; His_binding.				
DR	Pfam: PF02098; His_binding; 1.				
DR	ProDom: PD15245; His_binding; 1.				
DR	Signal: 3D-structure.				
FT	SIGNAL	1	19		
FT	CHAIN	20	190		
FT	DISULFID	67	188		
FT	DISULFID	138	167		
FT	TURN	23	24		
FT	HELI	27	30		
FT	HELI	31	33		
FT	HELI	36	41		
FT	TURN	42	45		
FT	STRAND	48	53		
FT	STRAND	57	58		
FT	TURN	59	61		
FT	STRAND	62	63		
FT	TURN	64	64		
FT	STRAND	66	76		

## FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN

```

FT TURN 77 80
FT STRAND 81 89
FT TURN 91 92
FT STRAND 97 106
FT TURN 109 110
FT STRAND 116 121
FT TURN 122 123
FT STRAND 126 135
FT TURN 136 137
FT STRAND 138 143
FT STRAND 152 157
FT TURN 160 161
FT HELIX 165 174
FT TURN 175 177
FT STRAND 181 182
FT TURN 186 188
SQ SEQUENCE 190 AA; 21464 MW; 6923A3E902552B6F CRC64;

Query Match 100.0%; Score 936; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.8e-74;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOPWMADEAANGAHQDAMKSLKADVENYVYVAVKATYKNDPVGNDFTCVGMADVNEDE 60
DB 20 NOPWMADEAANGAHQDAMKSLKADVENYVYVAVKATYKNDPVGNDFTCVGMADVNEDE 79
QY 61 KSIOAEFLFNNMADTNMQFATEKYAVKMGYNENAFRETEDEGQVFTVIAVSDNCD 120
DB 80 KSIOAEFLFNNMADTNMQFATEKYAVKMGYNENAFRETEDEGQVFTVIAVSDNCD 139
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNPANCLKFNENYAVGRETRVFTSACL 171
DB 140 VIYVPGTGDGNEGEYELMTTIDYDNPANCLKFNENYAVGRETRVFTSACL 190

RESULT 2
HBPM_RHIAP STANDARD; PRT; 190 AA.
ID HBPM_RHIAP STANDARD; PRT; 190 AA.
AC 077420;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Female-specific histamine-binding protein 1 precursor (FS-HBPI).
OS Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
ON NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=99288454; Pubmed=10360182;
RA Priesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT "Tick histamine-binding proteins: isolation, cloning, and three-
dimensional structure."
RL Mol. Cell 3:661-671(1999).
CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO
OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO
SUPPRESS INFLAMMATION DURING BLOOD FEEDING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U96080; AAC63106.1; -.
DR HSSP; 077421; 10FT.
DR InterPro; IPR002970; His_binding.
DR Pfam; PF02098; His_binding; 1.
DR ProDom; PD152455; His_binding; 1.

```

```

KW Signal.
FT SIGNAL. 1 18
FT CHAIN 19 190
FT DISULFID 66 187
FT DISULFID 137 166
SQ SEQUENCE 190 AA; 21370 MW; 855BE151A90053B1 CRC64;

Query Match 67.3%; Score 630; DB 1; Length 190;
Best Local Similarity 66.5%; Pred. No. 8.7e-48;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 NOPWMADEAANGAHQDAMKSLKADVENYVYVAVKATYKNDPVGNDFTCVGMADVNEDE 60
DB 19 DKPYMADEAANGAHQDAMKSLKADVENYVYVAVKATYKNDPVGNDFTCVGMADVNEDE 78
QY 61 KSIOAEFLFNNMADTNMQFATEKYAVKMGYNENAFRETEDEGQVFTVIAVSDNCD 120
DB 79 KNEVAMFPMFNNMADTVYQHTFEKATPDKMGYNENAFRETEDEGQVFTVIAVSDNCD 138
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNPANCLKFNENYAVGRETRVFTSACL 170
DB 139 VIYVPGTGDGNEGEYELMTTIDYDNPANCLKFNENYAVGRETRVFTSACL 188

RESULT 3
HBPM_RHIAP STANDARD; PRT; 200 AA.
ID HBPM_RHIAP STANDARD; PRT; 200 AA.
AC 077422;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Male-specific histamine-binding salivary protein precursor (MS-HBP).
OS Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
ON NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=99288454; Pubmed=10360182;
RA Priesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT "Tick histamine-binding proteins: isolation, cloning, and three-
dimensional structure."
RL Mol. Cell 3:661-671(1999).
CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO
OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO
SUPPRESS INFLAMMATION DURING BLOOD FEEDING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U96082; AAC63108.1; -.
DR HSSP; 077421; 10FT.
DR InterPro; IPR002970; His_binding.
DR Pfam; PF02098; His_binding; 1.
DR ProDom; PD152455; His_binding; 1.
FT SIGNAL.
FT CHAIN 1 18
FT DISULFID 65 193
FT DISULFID 137 169
SQ SEQUENCE 200 AA; 22851 MW; C46A1C8C6BCAA008 CRC64;

Query Match 35.4%; Score 331.5; DB 1; Length 200;
Best Local Similarity 42.0%; Pred. No. 6.4e-22;

```



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M17391; AAC63063.1; -

DR InterPro: IPR002479; CM\_binding.

DR InterPro: IPR003318; Glyco\_hydro\_70.

DR Pfam: PF01473; CM\_binding\_1; 16.

DR Pfam: PF02324; Glyco\_hydro\_70; 1.

KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.

FT SIGNAL 1 38

FT CHAIN 1 38

FT DOMAIN 39 1050 GLUCOSYLTRANSFERASE-I.

FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).

FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.

FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.

FT REPEAT 1163 1213 AC REPEAT.

FT REPEAT 1227 1277 AC REPEAT.

FT REPEAT 1292 1342 AC REPEAT.

FT REPEAT 1352 1399 AC REPEAT.

FT REPEAT 1406 1455 AC REPEAT.

FT REPEAT 1465 1512 AC REPEAT.

FT REPEAT 1519 1568 AC REPEAT.

FT REPEAT 1582 1597 AC REPEAT (INCOMPLETE).

FT SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64.

Query Match 9.78; Score 91; DB 1; Length 1597; Best Local Similarity 24.4%; Pred. No. 4.9; Matches 49; Conservative 21; Mismatches 55; Indels 76; Gaps 12;

QY 2 QPDM-----ADEANGAHQ-DAMKSLKADYENYVYV-----KATKK-N 38

DB 341 QPDMNGSEKRYDDHQLNGKLFKDNOSDLPPDQSNRYLRNRPPTNGSLSDFFTYNNAN 400

QY 39 DPVAGNDFTCGVYVANDVNEDEKSIQAEFL-----FMNADTNMOPATEKYTA 86

DB 401 DPLGVEL-----LLANDVDNSNPVQAEQLNMLHYLLNFGTYAKADADAN--FDSIRVDA 454

QY 87 YKMYGVYRENAFRTEDGVFTDYVLAISDNCNDVYVPTDGNESG-----YELMT- 138

DB 455 V-----DNVADDLQISSDYLLKAAV--GIDNNKNNANHVSIWEAMSD 495

QY 139 -----TDYDNIPANCKNFK 152

DB 496 NDFPYLHDDGDNL--MNNDNKF 515

RESULT 6

YL2\_SCHPO STANDARD; PRT; 542 AA.

AC 09P378; P78853;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein C19B12.02c in chromosome I precursor.

GN SPAC19B12.02C.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomyces.

OX NCBI\_TaxID=4896;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Soutouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Stelson J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tvey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,

RA Wellings J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rocher M., Galliard C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe."

RL Nature 415:671-680(2002).

RN [2]

RP SEQUENCE OF 268-542 FROM N.A.

RC STRAIN=PR745;

RX MEDLINE=98162722; PubMed=9501991;

RA Yoshioaka S., Kato K., Nakai K., Okayama H., Nojima H.;

RT "Identification of open reading frames in Schizosaccharomyces pombe

RT cDNAs."

RL DNA Res. 4:363-369(1997).

CC -1- SIMILARITY: Belongs to the GAS1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AL390814; GAC00550.1; -

DR GeneDB: D89203; BAI3864.1; -

DR InterPro: IPR004886; GAS1.

DR Pfam: PF03198; GAS1; 1.

KW Hypothetical protein; Glycoprotein; Signal.

FT SIGNAL 1 19

FT CHAIN 1 19

FT DOMAIN 434 520

FT CARBOHYD 35 35

FT CARBOHYD 91 91

FT CARBOHYD 161 161

FT CARBOHYD 249 249

FT CARBOHYD 279 279

FT CARBOHYD 406 406

FT CARBOHYD 484 484

FT CARBOHYD 502 502

FT CARBOHYD 509 509

FT CARBOHYD 520 520

FT CONFLICT 268 278

FT CONFLICT 285 293

FT SEQUENCE 542 AA; 58115 MW; 1D243FDDA5AD2BC3A CRC64;

Query Match 9.6%; Score 89.5; DB 1; Length 542; Best Local Similarity 25.1%; Pred. No. 1.8; Matches 44; Conservative 22; Mismatches 60; Indels 49; Gaps 10;

QY 13 AHODAMSLKADYENYVYVYKATYKNDPYVGNDF-----TCGVAMANDVN 57

DB 102 AFDQDAGTIVYSLDIAQY-----EAISSDPTWVDFSGRYEVVDSLAPYDNMGLFING--N 156

QY 58 EDEKSIQAEFLFMNNDTN--MOPATEKYVAVKMY-----GVNRENARRYTEGQVYTDVI 112

DB 157 E-----VIONNTNTNAAAFKAAVADVVSIISSG--RQIPYGYSTNDEEYTRDPM 206

QY 113 AY-----SDNCDVY-----YVPTDGNESGEYELMTDYDNIIPANCKNFKNEY 155

DB 207 AYTFDGGDDDDHDFGINTYEMCGSDSYSSGYQRTFEFSMTVPMI--FSEF 259





DR PRINTS: PR00736; GLHYDRLASE15.  
 DR PROSITE: PS00820; GLUCOAMYLASE; 1.  
 KW Hydrolyase; glycosidase; Polysaccharide degradation; glycoprotein;  
 RM Signal.  
 FT CHAIN 1 18 POTENTIAL.  
 FT ACT\_SITE 19 624 GLUCOAMYLASE.  
 FT ACT\_SITE 340 340 CATALYTIC BASE (BY SIMILARITY).  
 FT ACT\_SITE 343 343 GENERAL ACID CATALYST (BY SIMILARITY).  
 FT ACT\_SITE 344 344 INTERACT WITH SUBSTRATES (BY SIMILARITY).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 624 AA; 68980 MW; 74C2CDFB43FE71B4 CRC64;

Query Match 9.08; Score 84; DB 1; Length 624;  
 Best Local Similarity 21.58; Pred. No. 6.5;  
 Matches 45; Conservative 22; Mismatches 64; Indels 78; Gaps 9;

OY 10 ANGAHODA-----WKSLADY-----ENYYMKATYKN 38  
 DB 299 ANGSDQTVREQLYSDSKYQVAFIKDLQFVANSWSPSPDLMEERESAHFTYRVOR 358  
 OY 39 DPVWGNDFCTGVWANDVNEDEKSIQAEFLMNNADTNMOF-----79  
 DB 359 ALLIGADF-----ANDMGHLSDKLTKQASKSDLPFLPFWSARQLIYEGPVLRGK 412  
 OY 80 -ATEKTAIV--KMGYRENAFRETEBDGVFTDVIAYSDNCDVIYVPGTGNESGYEL 136  
 DB 413 YSKYKDISVYGVNMGVANDVFSY-TND-QILATAYGVSTFIDVYKVAWTTSDSESKPL 470  
 OY 137 WTTDYDIPANCINKNEVAVGRETRDVE 165  
 DB 471 -----GIP-----VGRYPEDVY 482

RESULT 9  
 CT99\_HUMAN  
 AC Q9N0D5; STANDARD; PRT; 404 AA.  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Hypothetical protein C20orf99.  
 GN C20orf99.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Cleeg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehesvistalo M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vautin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 "The DNA sequence and comparative analysis of human chromosome 20.";  
 Nature 414:865-871(2001).

-1- SIMILARITY: Contains 3 CCHC-type zinc fingers.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AL034548; CAB81631.1;  
 CC Genew: HGNC:16230; C20orf99.  
 DR InterPro: IPR001878; Znf.CCHC.  
 DR Pfam: PF000098; zfc-CCHC; 3.  
 DR PRINTS: PR00939; C2HC2NFINGER.  
 DR SMART: SM00343; Znf.C2HC; 2.  
 DR PROSITE: PS50158; ZF\_CCHC; 2.  
 KW Hypothetical protein; zinc-finger; Repeat.  
 FT ZN\_FING 334 351 CCHC-TYPE 1.  
 FT ZN\_FING 352 369 CCHC-TYPE 2.  
 FT ZN\_FING 371 388 CCHC-TYPE 3.  
 FT DOMAIN 67 71 POLY-GLY.  
 FT DOMAIN 126 134 POLY-ALA.  
 FT DOMAIN 155 159 POLY-ALA.  
 FT SEQUENCE 404 AA; 43618 MW; C3B728090DD27A4B CRC64;

Query Match 8.98; Score 83.5; DB 1; Length 404;  
 Best Local Similarity 21.88; Pred. No. 4.3;  
 Matches 37; Conservative 30; Mismatches 64; Indels 39; Gaps 8;

OY 2 QPDMADEAANGAHODAMKSLKADVENYYVAKATYKN-----PYWGNDFCTGVWANDVN 57  
 DB 138 RPEBAEDPAERPIQDEPAALAAAGPGKGRFLVRICFGDEGACPT--RDFVYVALLRISIG 195  
 OY 58 EDEKSIQAEFLFMNADTNMOF-ATEKVTA-VMYGYNREN-----96  
 DB 196 MDPSDIYAVIQIGSRFEDVFSRSAEKLAFLVYEKKRQEDQWENFVILGSKSISKT 255  
 OY 97 ---AFRETEBDGVFTDVIAYSDNCDVIYV--GTDGNEEGYELMTTDT 141  
 DB 256 LFTLFRENEVDVE---DIVTWLKRHCDVLAVPKVD---RGIMTGEY 298

RESULT 10  
 FBN2\_HUMAN  
 ID FBN2\_HUMAN STANDARD; PRT; 2911 AA.  
 AC P35556;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrillin 2 precursor.  
 GN FB2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94165150; PubMed=8120105;  
 RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sangulneti C.,  
 RA Bonadio J., Mecham R.P., Ramirez F.;  
 "Structure and expression of fibrillin-2, a novel microfibrillar





```
FT DISULFID 321 333 BY SIMILARITY.
FT DISULFID 328 342 BY SIMILARITY.
FT DISULFID 344 357 BY SIMILARITY.
FT DISULFID 497 509 BY SIMILARITY.
FT DISULFID 504 518 BY SIMILARITY.
FT DISULFID 520 532 BY SIMILARITY.
FT DISULFID 538 548 BY SIMILARITY.
FT DISULFID 543 557 BY SIMILARITY.
FT DISULFID 559 572 BY SIMILARITY.
FT DISULFID 578 590 BY SIMILARITY.
FT DISULFID 585 599 BY SIMILARITY.
FT DISULFID 601 614 BY SIMILARITY.
FT DISULFID 620 631 BY SIMILARITY.
FT DISULFID 626 640 BY SIMILARITY.
FT DISULFID 642 655 BY SIMILARITY.
FT DISULFID 661 672 BY SIMILARITY.
FT DISULFID 667 681 BY SIMILARITY.
FT DISULFID 683 696 BY SIMILARITY.
FT DISULFID 771 783 BY SIMILARITY.
FT DISULFID 778 792 BY SIMILARITY.
FT DISULFID 794 807 BY SIMILARITY.
FT DISULFID 813 825 BY SIMILARITY.
FT DISULFID 820 834 BY SIMILARITY.
FT DISULFID 836 849 BY SIMILARITY.
FT DISULFID 855 865 BY SIMILARITY.
FT DISULFID 860 874 BY SIMILARITY.
FT DISULFID 876 889 BY SIMILARITY.
FT DISULFID 958 970 BY SIMILARITY.
FT DISULFID 965 979 BY SIMILARITY.
FT DISULFID 981 994 BY SIMILARITY.
FT DISULFID 1076 1088 BY SIMILARITY.
FT DISULFID 1083 1097 BY SIMILARITY.
FT DISULFID 1099 1112 BY SIMILARITY.
FT DISULFID 1118 1130 BY SIMILARITY.
FT DISULFID 1125 1139 BY SIMILARITY.
FT DISULFID 1141 1155 BY SIMILARITY.
FT DISULFID 1161 1173 BY SIMILARITY.
```

Query Match 8.98; Score 83.5; DB 1; Length 2911;  
Best Local Similarity 29.0%; Pred. No. 44;  
Matches 49; Conservative 16; Mismatches 45; Indels 59; Gaps 13;

```
QY 7 DEANAGHODAMSLKADVENYVYKATYK---NDPVANGNDTCVGVANDVNEDEKSI 63
DB 1409 DECSNCTHO---CSINACVN---TPGSYRCACSGFTGDEGTC---SDVDECAENI 1456
QY 64 QAEEFLMNNADTMMOFATEKVTAVKMYGYNRENAFRETEDGOVFT---DVIAYS-DNC 119
DB 1457 -----NLCEMGQC-----LNVPGRVRCCECMG---FTPADSRSCQDIDEC 1494
QY 120 DV---IYVPGTDGN-----EEGYEL-----WTTDYDNI---PANCLN 150
DB 1495 SPTNLCVSGTCNNLPGMFHCICDDGYELDRGTGNCYDIDECADPINCYN 1543

RESULT 11
XRNA_RUMFL STANDARD; PRT; 954 AA.
ID XRNA_RUMFL
AC P29126;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bifunctional endo-1,4-beta-xylanase Xyla precursor (EC 3.2.1.8).
GN Xyna.
OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Ruminococcus.
OX NCBI_TaxID=1265;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17;
RX MEDLINE=92261318; PubMed=1584021;
RA Zhang J.-X., Flint H.J.;
```

"A bifunctional xylanase encoded by the xyna gene of the rumen cellulolytic bacterium Ruminococcus flavefaciens 17 comprises two dissimilar domains linked by an asparagine/glutamine-rich sequence." Mol. Microbiol. 6:1013-1023(1992).

-1- FUNCTION: XYLANASE DOMAIN 1 RELEASES MORE XYLO-OLIGOSACCHARIDES AND DOMAIN 2 MORE XYLOSE.

-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic linkages in xylans.

-1- PATHWAY: Xylan degradation.

-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY G (FAMILY 11 OF GLYCOSYL HYDROLASES).

-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL HYDROLASES).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
DR EMBL: Z11127; CAAT7476.1; -
DR PIR: S20907; S20907.
DR HSSP: P48793; 1XND.
DR InterPro: IPR001000; Glyco_hydro_10.
DR InterPro: IPR001137; Glyco_hydro_11.
DR Pfam: PF00031; Glyco_hydro_10; 1.
DR Pfam: PF00457; Glyco_hydro_11; 1.
DR PRINTS: PR00134; GLHTRILASE10.
DR PRINTS: PR00911; GLHYDRLASE11.
DR SMART: SM00633; Glyco_10; 1.
DR PROSITE: PS00591; GLYCOSYL_HYDROL_F10; 1.
DR PROSITE: PS00776; GLYCOSYL_HYDROL_F11; 1.
DR PROSITE: PS00777; GLYCOSYL_HYDROL_F12; 1.
KM xylan degradation; Hydrolase; Glycosidase; Multifunctional enzyme;
KW Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 954 OR 28, OR 29 (POTENTIAL).
FT DOMAIN 28 244 BIFUNCTIONAL ENDO-1,4-BETA-XYLANASE XYLA.
FT DOMAIN 245 622 XYLANASE DOMAIN 1.
FT DOMAIN 623 954 ASN/GLN/TRP-RICH (LINKER).
FT ACT_SITE 122 122 XYLANASE DOMAIN 2.
FT ACT_SITE 223 223 NOCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 774 774 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 884 884 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 884 884 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 954 AA; 111362 MW; 1033567DAB52EBD CRC64;
```

Query Match 8.88; Score 82.5; DB 1; Length 954;  
Best Local Similarity 24.38; Pred. No. 14;  
Matches 43; Conservative 26; Mismatches 37; Indels 71; Gaps 13;

```
QY 3 PDW---ADEANGAH--ODAW-KSLKADVENYVYKADYKNDPVANGNDTCVGVANDV 56
DB 722 PDWFFRENSQNGAIVSKIDIMNORLESMTKNTFALKSKSPINLDVSYD-VC----- 772
QY 57 NDEKSIQAEEFLMNNADTMMOFATEKVTAVKMYGYNR--ENAFRETEDGOVFTDVIAV 114
DB 773 NE-----LEFNNGG--GMRGA--DNSMWVKIYGDSFVIAFKVARQ-----Y 811
QY 115 SDDNCVITYVPGTDGNEBGEYELWTTDYDNIIPANC---LKNREYAVAGRTDVFYSA 168
DB 812 A-----PAGCKLYLNDYNEY-IPAKTNDIYNNNA 838

RESULT 12
RPA2_DROME STANDARD; PRT; 1129 AA.
ID RPA2_DROME
AC P20028;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DNA-directed RNA polymerase I 135 kDa polypeptide (EC 2.7.7.6)
```

```

DE (RNA polymerase I subunit 2).
GN Rpl135.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90158499; PubMed=2482932;
RA Klettermann R., Stizler S., Selfarth W., Petersen G., Bautz E.K.F.;
RT "Primary structure and functional aspects of the gene coding for the
RT second-largest subunit of RNA polymerase III of Drosophila.";
RL Mol. Gen. Genet. 219:373-380(1989).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES. RNA POLYMERASE I IS ESSENTIALLY USED TO TRANSCRIBE
CC RIBOSOMAL DNA UNITS.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14
CC DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS THE SECOND LARGEST
CC COMPONENT OF RNA POLYMERASE I.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X17298; CA35185.1; -.
DR PIR: J00354; J00354.
DR FlyBase; FBgn003278; Rpl135.
DR InterPro; IPR001572; RNA_pol_8.
DR Pfam; PF04563; RNA_pol_RpD2_1; 1.
DR Pfam; PF04561; RNA_pol_RpD2_2; 1.
DR Pfam; PF04565; RNA_pol_RpD2_3; 1.
DR Pfam; PF04567; RNA_pol_RpD2_5; 1.
DR Pfam; PF00562; RNA_pol_RpD2_6; 1.
DR Pfam; PF04560; RNA_pol_RpD2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; zinc;
KW Zinc-finger; Nuclear protein.
FT ZN.FING 1061 1093
FT SEQUENCE 1129 AA; 128414 MW; E0A15FE1BCEFB18D7 CRC64;
SQ

```

Query Match 8.7%; Score 81.5; DB 1; Length 1129;  
Best Local Similarity 25.0%; Pred. No. 22;  
Matches 21; Conservative 15; Mismatches 37; Indels 11; Gaps 3;

```

OY 11 NGAHQDAWKSLLK--DVENYVYVYKATYKNDP-----W---GNDFTCVGVANDVED 59
DB 158 HGEHDEWEGCIFVIRGNEKIVRMILMTRNHPICVKRSMWDRCGNFSDGLMLVQTVARE 217
OY 60 EKSTQAEFLFMNNADTNNOFATEK 83
DB 218 ESSLSNVVHYLNGTAKFMESHVK 241

```

RESULT 13  
CICA\_PAEPP STANDARD: PRT; 695 AA.  
AC P57092;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Parasporal crystal protein cry18Ca (Parasporal delta-endotoxin
DE Cryxviic(a)) (Crystalline parasporal protoxin) (78 kDa crystal
DE protein).
GN CRY18CA OR CRYXVIIC(A).
OS Paenibacillus popilliae (Bacillus popilliae).
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
OX NCBI_TaxID=78057;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 14706;
RA Patel R., Yousten A.A., Rippere K.;
RT "Detection of two new cry genes in Paenibacillus popilliae.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO THE BRUSH BORDER MEMBRANE VESICLES OF SCARAB
CC LARVAE AND DAMAGES THE GUT WALL, SOMEHOW TO ALLOW THE VEGETATIVE
CC CELLS OF P. POPILLIAE TO ENTER THE HEMOLYMPH (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF169251; AAF89668.1; -.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
FT TOXIN 695 AA; 78259 MW; 406AC9154D75E070 CRC64;
SQ

```

Query Match 8.5%; Score 80; DB 1; Length 695;  
Best Local Similarity 24.5%; Pred. No. 16;  
Matches 45; Conservative 13; Mismatches 68; Indels 58; Gaps 8;

```

OY 18 WKSLEADVENYVYVYKATYKNDPVGWANDFTCVGVANDVEDKSIQAEFLFMNNADTNM 77
DB 306 WSLK--YVNYVYSTANLYN-----IGDKNVNGEVSISWPFNFYIQKRS 351
OY 78 QFATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDNCDVIYYPGTDGNEGYELW 137
DB 352 NYVLGVSQYAM-RMSYTNPF-----FGYTIQDILYNTFASVIGGVNGPQIGQOLS 401
OY 138 TT-----DYNIIPANC-----LNKFNEY-----AVGERTR-D 163
DB 402 TTEDQLVQOQARADIPVDQIPINCLNRPLEVYATRFNELTSLGTAGVGGFVRSD 461
OY 164 VFTS 167
DB 462 VFIS 465

```

RESULT 14  
FBN2\_MOUSE STANDARD: PRT; 2907 AA.  
ID FBN2\_MOUSE  
AC 061555; 063957;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrillin 2 precursor.  
GN FBN2 OR FBN-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]



```
FT DISULFID 975 988 BY SIMILARITY.
FT DISULFID 1070 1082 BY SIMILARITY.
FT DISULFID 1077 1091 BY SIMILARITY.
FT DISULFID 1093 1106 BY SIMILARITY.
FT DISULFID 1112 1124 BY SIMILARITY.
FT DISULFID 1119 1133 BY SIMILARITY.
FT DISULFID 1135 1149 BY SIMILARITY.
FT DISULFID 1155 1167 BY SIMILARITY.
FT DISULFID 1162 1176 BY SIMILARITY.
FT DISULFID 1178 1191 BY SIMILARITY.
FT DISULFID 1197 1209 BY SIMILARITY.
FT DISULFID 1204 1218 BY SIMILARITY.
FT DISULFID 1220 1233 BY SIMILARITY.
FT DISULFID 1239 1250 BY SIMILARITY.
FT DISULFID 1246 1259 BY SIMILARITY.
FT DISULFID 1261 1274 BY SIMILARITY.
FT DISULFID 1280 1292 BY SIMILARITY.
FT DISULFID 1287 1301 BY SIMILARITY.
FT DISULFID 1303 1316 BY SIMILARITY.
FT DISULFID 1322 1334 BY SIMILARITY.
FT DISULFID 1329 1343 BY SIMILARITY.
FT DISULFID 1345 1358 BY SIMILARITY.
FT DISULFID 1364 1377 BY SIMILARITY.
FT DISULFID 1371 1386 BY SIMILARITY.
FT DISULFID 1388 1399 BY SIMILARITY.
FT DISULFID 1405 1418 BY SIMILARITY.
FT DISULFID 1412 1427 BY SIMILARITY.
FT DISULFID 1429 1440 BY SIMILARITY.
FT DISULFID 1446 1458 BY SIMILARITY.
FT DISULFID 1453 1467 BY SIMILARITY.
FT DISULFID 1469 1482 BY SIMILARITY.
FT DISULFID 1488 1499 BY SIMILARITY.
FT DISULFID 1494 1508 BY SIMILARITY.
FT DISULFID 1510 1523 BY SIMILARITY.
FT DISULFID 1529 1540 BY SIMILARITY.
FT DISULFID 1535 1549 BY SIMILARITY.
FT DISULFID 1551 1564 BY SIMILARITY.
FT DISULFID 1647 1659 BY SIMILARITY.
FT DISULFID 1654 1668 BY SIMILARITY.
FT DISULFID 1670 1683 BY SIMILARITY.
FT DISULFID 1689 1701 BY SIMILARITY.
FT DISULFID 1696 1710 BY SIMILARITY.
```

Query Match 8.5%; Score 79.5; DB 1; Length 2907;  
Best Local Similarity 29.0%; Pred. No. 98;  
Matches 49; Conservative 15; Mismatches 46; Indels 59; Gaps 13;

```
Oy 7 DEANGAHQDAWKSLLADVENVYVAVATYK--NDPYGNDFTCVGYMANDVNEDEKSI 63
    |||||
Db 1403 DEANGTHQ---CSINACVN---TPGSYRCACSEGFTGDFGFC-----SDVDECAEN- 1449
Oy 64 QAEFLFMNADTNMGFAETKTAVKMGVYRENAFRYTEDEGQVFT---DVIAVSD-DNC 119
    |-----TNLCENOC-----LNVPAGVRCCEMG--FTPASDSRSCQDIDEC 1488
Db 1450 -----TNLCENOC-----LNVPAGVRCCEMG--FTPASDSRSCQDIDEC 1488
Oy 120 DV--IYVPGTDGN-----EEGYEL-----MTTDVNT--PANCIN 150
    |||||
Db 1489 SFQNICVPGTCNNLPGMFHCICDDGIELDRTGCTDDECADPINCYN 1537
```

RESULT 15  
ID ITB4\_HUMAN STANDARD; PRT: 1822 AA.  
AC P16144; O14690; O14691; O15339; O15340; O15341; Q9U104;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Integrin beta-4 precursor (GP150) (CD104 antigen).  
GN ITGB4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OC NCBI\_TaxID=9606;

```
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM BETA-4A).  
RX MEDLINE-90183973; PubMed-2311577;  
RA Suzuki S., Naitoh Y.;  
RT "Amino acid sequence of a novel integrin beta 4 subunit and primary  
RN expression of the mRNA in epithelial cells.";  
RL EMBO J. 9:757-763(1990).  
RZ [2]  
RP SEQUENCE FROM N.A. (ISOFORM BETA-4B).  
RX MEDLINE-90183974; PubMed-2311578;  
RA Hogervorst F., Kulkman I., von Dem Borne A.E.G.K., Sonnenberg A.;  
RT "Cloning and sequence analysis of beta-4 cDNA: an integrin subunit  
RN that contains a unique 118 kd cytoplasmic domain.";  
RL EMBO J. 9:765-770(1990).  
RZ [3]  
RP SEQUENCE FROM N.A. (ISOFORM BETA-4C).  
RC TISSUE-Pancreas;  
RX MEDLINE-91009492; PubMed-1976638;  
RA Tamura R.N., Rozzo C., Starr L., Chambers J., Reichardt L.F.,  
RN Cooper H.M., Quaranta V.;  
RT "Epithelial integrin alpha 6 beta 4: complete primary structure of  
RN alpha 6 and variant forms of beta 4.";  
RL J. Cell Biol. 111:1593-1604(1990).  
RZ [4]  
RP SEQUENCE FROM N.A. (ISOFORMS BETA-4A; BETA-4B AND BETA-4C).  
RX MEDLINE-97338298; PubMed-9194858;  
RA Pulkkinen L., Kurtz K.S., Xu Y., Bruckner-Tuderman L., Uitto J.;  
RT "Genomic organization of the integrin beta 4 gene (ITGB4): a  
RN homozygous splice-site mutation in a patient with junctional  
RT epidermolysis bullosa associated with pyloric atresia.";  
RL Lab. Invest. 76:823-833(1997).  
RZ [5]  
RP SEQUENCE FROM N.A. (ISOFORMS BETA-4A; BETA-4B AND BETA-4C).  
RC TISSUE-Lung;  
RX MEDLINE-97311186; PubMed-9166594;  
RA Iacovacci S., Gagnoux-Palacios L., Zambruno G., Meneguzzi G.,  
RN D'Alessio M.;  
RT "Genomic organization of the human integrin beta 4 gene.";  
RL Mamm. Genome 8:448-450(1997).  
RZ [6]  
RP REVISIONS.  
RA D'Alessio M.;  
RN submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
RZ [7]  
RP SEQUENCE FROM N.A. (ISOFORM BETA-4E).  
RX MEDLINE-97350870; PubMed-9207246;  
RA van Leusden M.R., Kulkman I., Sonnenberg A.;  
RT "The unique cytoplasmic domain of the human integrin variant beta4E is  
RN produced by partial retention of intronic sequences.";  
RL Biochem. Biophys. Res. Commun. 235:826-830(1997).  
RZ [8]  
RP SEQUENCE OF 28-46.  
RX MEDLINE-89251596; PubMed-2542022;  
RA Kajlji S., Tamura R.N., Quaranta V.;  
RT "A novel integrin (alpha E beta 4) from human epithelial cells  
RN suggests a fourth family of Integrin adhesion receptors.";  
RL EMBO J. 8:673-680(1989).  
RZ [9]  
RP ALTERNATIVE SPLICING (ISOFORM BETA-4D).  
RX MEDLINE-95073005; PubMed-7982032;  
RA Clarke A.S., Lotz M.M., Mercutio A.M.;  
RT "A novel structural variant of the human beta 4 integrin cDNA.";  
RN Cell Adhes. Commun. 2:1-6(1994).  
RZ [10]  
RP VARIANTS EB-PA TYR-61, CYS-252; ARG-562 AND TRP-1281.  
RX MEDLINE-99011257; PubMed-9792864;  
RA Pulkkinen L., Rouan F., Bruckner-Tuderman L., Wallerstein R.,  
RN Garzon M., Brown T., Smith L., Carter W.G., Uitto J.;  
RT "Novel ITGB4 mutations in lethal and nonlethal variants of  
RN epidermolysis bullosa with pyloric atresia: missense versus  
RT nonsense.";  
RL Am. J. Hum. Genet. 63:1376-1387(1998).  
RZ [11]
```

RP VARIANT EB-PA GUY-245.  
 RX MEDLINE-98082956; PubMed-9422533;  
 RA Pulkkinen L., Kim D.U., Uitto J.;  
 RT "Epidermolysis bullosa with pyloric atresia: novel mutations in the  
 RT beta-4 integrin gene (ITGB4)."  
 RL Am. J. Pathol. 152:157-166(1998).  
 RN [12]  
 RP VARIANT EB-PA PRO-156.  
 RX MEDLINE-98206430; PubMed-9546354;  
 RA Pulkkinen L., Bruckner-Tuderman L., August C., Uitto J.;  
 RT "Compound heterozygosity for missense (L156P) and nonsense (R554X)  
 RT mutations in the beta-4 integrin gene (ITGB4) underlies mild,  
 RT nonlethal phenotype of epidermolysis bullosa with pyloric atresia."  
 RL Am. J. Pathol. 152:933-941(1998).  
 RN [13]  
 RP VARIANTS EB-PA ARG-38.  
 RX MEDLINE-9911354; PubMed-9892956;  
 RA Mellierio J.E., Pulkkinen L., McMillan J.R., Lake B.D., Horn H.M.,  
 RA Tidman M.J., Harper J.I., McGrath J.A., Uitto J., Eady R.A.J.;  
 RT "Pyloric atresia-junctional epidermolysis bullosa syndrome: mutations  
 RT in the integrin beta4 gene (ITGB4) in two unrelated patients with  
 RT mild disease."  
 RL Br. J. Dermatol. 139:862-871(1998).  
 RN [14]  
 RP VARIANT EB-PA TRP-1281.  
 RX MEDLINE-20334107; PubMed-10873890;  
 RA Kambham N., Tanji N., Seigle R.L., Markowitz G.S., Pulkkinen L.,  
 RA Uitto J., D'Agati V.D.;  
 RT "Congenital focal segmental glomerulosclerosis associated with beta4  
 RT integrin mutation and epidermolysis bullosa."  
 RL Am. J. Kidney Dis. 36:190-196(2000).  
 RN [15]  
 RP VARIANT GABEB ASP-931.  
 RX MEDLINE-20235353; PubMed-10792571;  
 RA Inoue M., Tamai K., Shimizu H., Owaribe K., Nakama T., Hashimoto T.,  
 RA McGrath J.A.;  
 RT "A homozygous missense mutation in the cytoplasmic tail of beta4  
 RT integrin, G931D, that disrupts hemidesmosome assembly and underlies  
 RT non-herlitz junctional epidermolysis bullosa without pyloric  
 RT atresia."  
 RL J. Invest. Dermatol. 114:1061-1064(2000).  
 RN [16]  
 RP VARIANTS EB-PA.  
 RX MEDLINE-21149723; PubMed-11251584;  
 RA Ashton G.H.S., Sorelli P., Mellierio J.E., Keane F.M., Eady R.A.J.,  
 RA McGrath J.A.;  
 RT "Alpha 6 beta 4 integrin abnormalities in junctional epidermolysis  
 RT bullosa with pyloric atresia."  
 RL Br. J. Dermatol. 144:408-414(2001).  
 RN [17]  
 RP VARIANTS HIS-98 AND LEU-844.  
 RX MEDLINE-21183353; PubMed-11289717;  
 RA Hirano A., Nagai H., Harada H., Terada Y., Haga S., Kajiwara T.,  
 RA Emi M.;  
 RT "Nine novel single-nucleotide polymorphisms in the integrin beta4  
 RT (ITGB4) gene in the Japanese population."  
 RL J. Hum. Genet. 46:35-37(2001).  
 RN [18]  
 RP VARIANTS EB-PA TYR-131; CYS-252; ASP-273; CYS-283; ASP-325; PRO-336  
 RP AND HIS-1225, AND VARIANT GLN-1216.  
 RX MEDLINE-21226999; PubMed-11328943;  
 RA Nakano A., Pulkkinen L., Murrell D., Ritco J., Lucky A.W., Garzon M.,  
 RA Stevens C.A., Robertson S., Pfendner E., Uitto J.;  
 RT "Epidermolysis bullosa with congenital pyloric atresia: novel  
 RT mutations in the beta 4 integrin gene (ITGB4) and genotype/phenotype  
 RT correlations."  
 RL Pediatr. Res. 49:618-626(2001).  
 CC -1- FUNCTION: INTEGRIN ALPHA-6/BETA-4 IS A RECEPTOR FOR LAMININ. IT  
 CC PLAYS A CRITICAL STRUCTURAL ROLE IN THE HEMIDESMOSOME OF  
 CC EPITHELIAL CELLS.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-4  
 CC ASSOCIATES WITH ALPHA-6.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms-5;  
 CC Name-Beta-4C;  
 CC IsoId-P16144-1; Sequence-Displayed;  
 CC Name-Beta-4A;  
 CC IsoId-P16144-2; Sequence-VSP\_002749;  
 CC Name-Beta-4B;  
 CC IsoId-P16144-3; Sequence-VSP\_002749, VSP\_002750;  
 CC Name-Beta-4D;  
 CC IsoId-P16144-4; Sequence-VSP\_002749, VSP\_002751;  
 CC Name-Beta-4E;  
 CC IsoId-P16144-5; Sequence-VSP\_002747, VSP\_002748;  
 CC -1- TISSUE SPECIFICITY: INTEGRIN ALPHA-6/BETA-4 IS PREDOMINANTLY  
 CC EXPRESSED BY EPITHELIA. ISOFORM BETA-4E IS ALSO EXPRESSED IN COLON  
 CC AND PLACENTA. ISOFORM BETA-4E IS ALSO EXPRESSED IN EPIDERMIS,  
 CC LUNG, DUODENUM, HEART, SPLEEN AND STOMACH.  
 CC -1- DOMAIN: THE FIBRONECTIN TYPE III-LIKE DOMAINS BIND BPAG1 AND  
 CC PLECTIN AND PROBABLY ALSO RECRUIT BP230.  
 CC -1- DISEASE: DEFECTS IN ITGB4 ARE A CAUSE OF EPIDERMOLYSIS BULLOSA  
 CC LETALIS WITH PYLORIC ATRESIA (EB-PA), ALSO KNOWN AS JUNCTIONAL  
 CC EPIDERMOLYSIS BULLOSA WITH PYLORIC ATRESIA (PA-JEB) OR APIASIA  
 CC CUTIS CONGENITA WITH GASTROINTESTINAL ATRESIA. IT IS CHARACTERIZED  
 CC BY MUCCUTANEOUS FRAGILITY AND GASTROINTESTINAL ATRESIA, WHICH  
 CC MOST COMMONLY AFFECTS THE PYLORUS.  
 CC -1- DISEASE: DEFECTS IN ITGB4 ARE A CAUSE OF GENERALIZED ATROPHIC  
 CC BENIGN EPIDERMOLYSIS BULLOSA (GABEB). THIS NONLETHAL FORM OF  
 CC JUNCTIONAL EPIDERMOLYSIS BULLOSA IS CHARACTERIZED BY LIFE-LONG  
 CC BLISTERING OF THE SKIN, ASSOCIATED WITH HAIR AND TOOTH  
 CC ABNORMALITIES.  
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.  
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.  
 CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 1414  
 CC TO 1429 DUE TO A FRAMESHIFT.  
 CC -1- DATABASE: NAME-PROV; NOTE-CD guide CD104 entry;  
 CC WWW-http://www.ncbi.nlm.nih.gov/prov/cd/cd104.htm".  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

Query Match 8.48; Score 78.5; DB 1; Length 1822;  
 Best Local Similarity 23.18; Pred. No. 69;  
 Matches 28; Conservative 20; Mismatches 50; Indels 23; Gaps 5;  
 OY 42 MGN-----DPTGVNANDVNEDEKSIQAEFLFMN-----NADTMWOFAT-----EKV 84  
 DB 194 WPNSDPPEFSKKNVISTEDYDEFRNKLQGERISGNLDAPEGCDAILQTVCTRIDGMRP 253  
 OY 85 TAVKMGYNRENARFRETEDGQVFTDVIVAYSDNCDVIVYPTGNGEYELMTTDYDNI 144  
 DB 254 DSHLLLVFESFAFHLEAGANVLACIMSNDRCHL-----DPTGYTYGR--TQYPSV 307  
 OY 145 P 145  
 DB 308 P 308

Search completed: August 1, 2003, 12:15:51  
 Job time : 13.4229 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 12.7029 Seconds  
(Without alignments)  
1294.578 million cell updates/sec

Title: US-10-087-195-7  
Perfect score: 936  
Sequence: 1 NOPDMADEANGAHQDAWKS.....FNEYAVGRETRDYFTSACLE 171

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

1: PIR\_76: \*  
2: plr1: \*  
3: plr2: \*  
4: plr3: \*  
4: plr4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	11.2	1592	2 A38175	glucosyltransferase
2	91	9.7	450	2 H82891	hypothetical prote
3	88	9.4	606	2 H97012	hypothetical prote
4	86.5	9.2	254	2 T18282	hypothetical prote
5	85.5	9.1	1237	2 AC1583	internalin protein
6	85.5	9.1	1612	2 AB1347	probable peptidogl
7	84.5	9.0	858	1 A42239	adenylate cyclase
8	83.5	8.9	2918	2 A54105	fibrillin-2 precur
9	83	8.9	601	2 S57962	cspc protein - Clo
10	82.5	8.8	954	1 S20907	endo-1,4-beta-xyla
11	81.5	8.7	1129	2 J00354	DNA-directed transc
12	80	8.5	166	2 T51637	myb-related transc
13	80	8.5	249	2 E84717	probable MYB famil
14	79.5	8.5	362	2 A12159	hypothetical prote
15	79.5	8.5	1377	2 D90358	hypothetical prote
16	79.5	8.5	2907	2 A57278	fibrillin-2 precur
17	79	8.4	327	2 D89931	branched-chain alp
18	79	8.4	1872	2 T30888	vitellogenin - Ath
19	78.5	8.4	531	2 A84471	En/Spm-like transp
20	78.5	8.4	785	2 H90547	lipoprotein (impor
21	78.5	8.4	964	2 UC5345	integrin beta-4 pr
22	78.5	8.4	1875	2 A36429	L-lactate dehydrog
23	78	8.3	317	2 A11206	hypothetical prote
24	78	8.3	338	2 C86775	hypothetical prote
25	78	8.3	363	2 T20745	hypothetical prote
26	78	8.3	446	2 G84262	prophage p13 prote
27	78	8.3	595	2 B86788	glutamate dehydrog
28	77.5	8.3	421	1 A38158	hexon protein - hu
29	77.5	8.3	512	2 S37279	

30	77.5	8.3	1026	2 A49750	beta-galactosidase
31	77.5	8.3	1659	2 H97926	hypothetical prote
32	77.5	8.3	1659	2 G95057	endo-beta-N-acetyl
33	77.5	8.2	150	2 A59103	hypothetical prote
34	77	8.2	236	2 E89769	hypothetical prote
35	77	8.2	335	2 G64385	hypothetical prote
36	77	8.2	366	2 F97343	uncharacterized co
37	77	8.2	393	2 S06256	hypothetical prote
38	77	8.2	2893	2 A64556	toxin-like outer m
39	76.5	8.2	669	2 E71610	hypothetical prote
40	76.5	8.2	1219	2 H84464	probable helicase
41	76	8.1	153	2 T06182	reverse transcript
42	76	8.1	368	1 Q0BEH2	early nuclear anti
43	76	8.1	403	2 T09322	DNA polymerase pro
44	76	8.1	466	2 G72603	nitrate reductase
45	76	8.1	515	2 H72455	hypothetical prote

## ALIGNMENTS

## RESULT 1

A38175 glucosyltransferase precursor - Streptococcus sobrinus

C:Species: Streptococcus sobrinus

C>Date: 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change 15-Oct-1999

C:Accession: A38175

R:Abu, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.

J. Bacteriol. 173, 989-996, 1991

A:Title: Peptide sequences for sucrose splitting and glucan binding with in Streptococ

A:Reference number: A38175; MUID:91123227; PMID:1704006

A:Accession: A38175

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1592 <ABO>

A:Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g21703

C:Superfamily: cpl repeat homology

F:1093-1112/Domain: cpl repeat homology <CP1>

F:1222-1241/Domain: cpl repeat homology <CP2>

F:1287-1306/Domain: cpl repeat homology <CP3>

F:1350-1351/Domain: cpl repeat homology <CP4>

F:1352-1371/Domain: cpl repeat homology <CP5>

F:1402-1420/Domain: cpl repeat homology <CP6>

F:1465-1484/Domain: cpl repeat homology <CP7>

F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 11.2%; Score 105; DB 2; length 1592;

Best Local Similarity 25.4%; Pred. No. 0.52;

Matches 51; Conservative 19; Mismatches 55; Indels 76; Gaps 12;

Oy	2	OPDW-----ADEANGAHQ-DAMKSLKADVENYVM-----KATYK-N 38
Db	335	QPQNGSESEKPYDHLONGALLFDNGTDLPPDTQSNRYLNKRTPTNQTGSLDSRFTYNNP 394
Oy	39	DPVWGNDFTCGVANVDNEDEKSIQAEFL-----FNNNADTNQFATEKVTA 86
Db	395	DFLGGYDF-----LLANDVDNSNPVYQAEQLWMLYTLNFGSIYANDADANF----- 441
Oy	87	VKMTGVYRENAFRYETEDGVFTDVIVASDDNCVIVYGTGMEGC-----YELWT- 138
Db	442	-----DSIRVDAED-INVDAQDLQISSDYLAAY--GIDKNNKNNNNHVSIVEAMSD 489
Oy	139	-----TDVDMIPANCLNKF 152
Db	490	NDFPYLHDDGDL-MNMDNKF 509

## RESULT 2

H82891 hypothetical protein U0441 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: H82891



R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
Submitted to Genbank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mlt  
A:Reference number: A82870  
A:Accession: H82891  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-450 <GLA>  
A:Cross-references: GB:AE002140; GB:AF222894; NID:96899420; PIDN:AAF30853.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UN441  
A:Genetic code: SGC3

Query Match 9.7%; Score 91; DB 2; Length 450;  
Best Local Similarity 23.4%; Pred. No. 1.9;  
Matches 37; Conservative 23; Mismatches 60; Indels 38; Gaps 8;

Db 7 DEANAGHDAKSLADVENYVYKATYKNDPVGNDPTCVG---VANDVNEDEKS 62  
19 DENSESHIKIMKELMVLQNDVKIINHSYGWSP---NDPLNIGFYNFINGKEIN-DDKD 74  
QY 63 IQAEFL-----FMNADTNMOPATEKVTAVKMY-----GYNENAFRYE 101  
Db 75 FKRELLKYYIDLTKNDQIVNNTINKYFKKNTYFISYPIITNIFSAAGNSTD--IRTR 132  
QY 102 TEDQV---FTDVIAVS---DDNCVYVPGTDGNEE 132  
Db 133 NKNGHILEIMFIDISKYSFMDREFNONSIVYSTKNNE 170

RESULT 3  
H97012  
hypothetical protein CAC0915 [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: H97012  
; Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: H97012  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-606 <RUR>  
A:Cross-references: GB:AE001437; PIDN:AAK78891.1; PID:q15023815; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC0915

Query Match 9.4%; Score 88; DB 2; Length 606;  
Best Local Similarity 25.5%; Pred. No. 5;  
Matches 27; Conservative 15; Mismatches 34; Indels 30; Gaps 5;

QY 55 DVNEDKSKIQAEFLFMNADTNMOPATEKVTAVKMGYNENAFRYETEDGQVFTDVIAV 114  
Db 40 DTNAERQAI-SPTIY---GTNODFSNAKYTA-RRIGGRSTGYNNENNDNSA----- 86  
QY 115 SDNDCVIVYPGTDGNEGYELMTPTDYD-----NIPANCLNKENE 154  
Db 87 -----GTDMKNESDNTWLTLYDPKEXYNEPASVYTAFFHD 121

RESULT 4  
T18282  
hypothetical protein G6 - slime mold (Dictyostelium discoideum) plasmid  
C:Species: Dictyostelium discoideum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T18282; S28718  
R;Rieben Jr., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hugh  
Genetics 148, 1117-1125, 1998  
A:Title: Dictyostelium discoideum nuclear plasmid Ddp5 is a chimera related to the Ddp1

A:Reference number: 214684; MUID:98198836; PMID:9539429  
A:Accession: T18282  
A>Status: preliminary; translated from GB/EMBL/DDbJ  
A:Molecule type: DNA  
A:Residues: 1-254 <RIE>  
A:Cross-references: EMBL:000796; NID:92702254; PID:92702257; PIDN:AAC18633.1  
A:Experimental source: Plasmid Ddp5  
R;Gurniak, C.B.; Bang, A.G.; Noegel, A.A.  
Curr. Genet. 17, 321-325, 1990  
A:Title: Transcript and sequence analysis of a 5.1 kb contiguous fragment of Dictyost  
A:Reference number: S28718; MUID:90254839; PMID:2340592  
A:Accession: S28718  
A:Molecule type: DNA  
A:Residues: 37-254 <GUR>  
A:Cross-references: EMBL:X53237; NID:97280; PIDN:CAA37327.1; PID:97281  
A:Experimental source: Plasmid Ddp1  
C:Genetics:  
A:Gene: d-5  
A:Genome: plasmid

Query Match 9.2%; Score 86.5; DB 2; Length 254;  
Best Local Similarity 23.5%; Pred. No. 2.3;  
Matches 31; Conservative 12; Mismatches 48; Indels 41; Gaps 5;

QY 26 ENYVYKATYKNDPVGNDFTCVGVANDVNEDEKSIOAEFLFMNADTNMOPATEKYT 85  
Db 155 ENYVHTNITIOER-----SNRFCTDPLCHCKNE-----NIQNIIDFKTKCT 198  
QY 86 AVKMGYGNRNFARFETEDGQVFTDVIAVSDNDCVIVYPGTDGNEGYELMTPTDNDIP 145  
Db 199 P-----KYGASDSEPLS-----TLYNPKLDSNGMSEKSVTOEKNTS 235

QY 146 ANCLNKENEYAV 157  
Db 236 NNL-KINITYLI 245

RESULT 5  
AC1583  
internalin protein (LPXNG motif) homolog lln1204 [imported] - Listeria innocua (strai  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AC1583  
R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maltournam, A.;  
Ok, C.; Schlueter, T.; Simoes, N.; Tilleret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1583  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1237 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC96435.1; PID:q16413678; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lln1204

Query Match 9.1%; Score 85.5; DB 2; Length 1237;  
Best Local Similarity 27.3%; Pred. No. 21;  
Matches 41; Conservative 18; Mismatches 70; Indels 21; Gaps 7;

QY 21 LKADVENYVYMK-----TYKNDPVGNDFTCVGVANDVNEDEKSIOAEFL--FMN 72  
Db 788 IKATPKEVIVYDAVAGANITVYKEDSECKLAENSLTGNVGEETSSAKELGTTLRE 847  
QY 73 ADTNMQ--FATEKVTAVKMGYN---RENARFETEDGQVFTDVIAVSDNDCVIVYPG 126  
Db 848 MFSMNGEESLEEQVITYYSKNPVPKAKITQVYTDDEGMELAPETLS-GAVDEVYVA- 905  
QY 127 TDGNEGYELMTPTDYDNIIPANCLNKENEYA 156



Db 906 TAKFTGTGELIET-----PSNAEGKFESENA 930

## RESULT 6

AB1347

probable peptidoglycan bound protein (LPXTG motif) lmo2178 [imported] - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AB1347  
R:Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Science 294, 849-852, 2001  
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Meek, C.; Schluteter, T.; Simoes, N.; Tlerrrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1347  
A:Status: preliminary  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1612 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAD00256.1; PID:g16411648; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo2178

Query Match 9.18; Score 85.5; DB 2; Length 1612;  
Best Local Similarity 26.08; Pred. No. 29;  
Matches 44; Conservative 22; Mismatches 68; Indels 35; Gaps 9;

Db 4 DMADEA--ANGAHQADANKSLKADYENYVYWKATYKNDPVWGNFTCVGVANDVNEDEK 61  
513 DMTSATSANENFMUKNTNTNAYEITYSKIDFSRKIKNEIT-----DENGVASDAT 568

Oy 62 -SIAEFL-----FMNADTNMOKATEKVA---VKMYGNRENAFYETEDGQVF 108  
Db 569 IAIPDLKKEAGTIDYFNNT-----MTWKITANSRDKIKGNINITEFSTGVKALKSY 622

Oy 109 TDVAYSDDNCDVIVPGTQNEGSELMYTDYDNIIPAN-CLNKFNENA 156  
Db 623 T-VNAVYDNTNSVLTTEGKD-----YTIDKVTTPAGFYIOLIGDYA 662

## RESULT 7

A42239

adenylate cyclase (EC 4.6.1.1) germination stage - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A42239  
R:Plitt, G.S.; Milona, N.; Borleis, J.; Lin, K.C.; Reed, R.R.; Devreotes, P.N.  
Cell 69, 305-315, 1992  
A:Title: Structurally distinct and stage-specific adenylyl cyclase genes play different  
A:Reference number: A42239; MUID:92233467; PMID:1348970  
A:Accession: A42239  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A:Molecule type: DNA  
A:Residues: 1-858 <PTT>  
A:Cross-references: GB:M87278  
C:Superfamily: slime mold germination stage adenylyl cyclase; guanylate cyclase catalyt  
C:Keywords: phosphorus-oxygen lyase  
F:351-574/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 9.0%; Score 84.5; DB 1; Length 858;  
Best Local Similarity 23.28; Pred. No. 16;  
Matches 39; Conservative 31; Mismatches 61; Indels 37; Gaps 9;

Oy 5 WADENANGAHQADANKSLKADYENY-----VYWKATYKNDPVWG-NDFTCVGYNA 53  
Db 43 YSEENSNIOQMDLEKSSKQIITHNQNMAYLLSSIDLKALYYVNPNDRDNFNF----- 98  
Oy 54 NDVNEDEKSIQAEFLF-----MNAADTNM---QFATEKVTAVKMGVYNE-NAFRETEDG 105

Db 99 --LNTLTKNSEFOYLFIRKIKNNNDNRNCFEERKIKDFQIYSFDENTNSIHVANKS 156  
Oy 106 QVFTDVAYSDDNCDVIVPGTQNEGSELMYTDYDNI--IPANCLNK 151  
Db 157 SYFPIIAFPDINKDIIT-----GLDINSTDYANETIKKSIPIK 194

## RESULT 8

A54105

fibrillin-2 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002  
C:Accession: A54105; S17063; S31101  
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sangunetti, C.; Bonadio, J.; Mecha  
J. Cell Biol. 124, 855-863, 1994  
A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component pr  
A:Reference number: A54105; MUID:94165150; PMID:8120105  
A:Accession: A54105  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-2918 <ZHA>  
A:Cross-references: GB:U03272  
R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattel, M.G.; Sarfarazi, M.; Tsipouras,  
Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17063

A:Molecule type: mRNA  
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>  
A:Cross-references: EMBL:X62009  
R:Miliewicz, D.M.  
submitted to the EMBL Data Library, December 1992  
A:Reference number: S31101  
A:Accession: S31101

A:Molecule type: mRNA  
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P',  
A:Cross-references: EMBL:X62009  
C:Genetics:  
A:Gene: GDB:FBN2

A:Cross-references: GDB:128122; OMIM:121050  
A:Map position: 5q23-5q31  
C:Superfamily: fibrillin 1; EGF homology  
C:Keywords: extracellular protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-2918/Product: fibrillin-2 #status predicted <MNT>  
F:1245-1280/Domain: EGF homology <EGF>  
F:1970-2013/Domain: EGF homology <EGF>

Query Match 8.9%; Score 83.5; DB 2; Length 2918;  
Best Local Similarity 29.0%; Pred. No. 95;  
Matches 49; Conservative 16; Mismatches 45; Indels 59; Gaps 13;

Oy 7 DEANGAHQADANKSLKADYENYVYWKATYK---NDPVWGNFTCVGVANDVNEDEKSI 63  
Db 1409 DECSNGTHQ---CSIMACVN---TPESYRCACSEGTGGGFC-----SDVECAENI 1456

Oy 64 QAEFLNNADTNMOKATEKVAVKMGYNRENAFYETEDGQVFT---DVAYSDD-ENC 119  
Db 1457 -----NLCEKGC-----LWPGVATRCCEMG---FTPASDSRSCODIDEC 1494

Oy 120 DV--IYVPGTGN-----ERGYEL-----WTTDYDNI--PANCLN 150  
Db 1495 SFQNVCSGTCNNLPGMFHCICDDGYELDRGTGACTDIDECADPINCYN 1543

## RESULT 9

S57962

cspc protein - Clostridium acetobutylicum (fragment)  
C:Species: Clostridium acetobutylicum  
C:Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 23-Mar-2001  
C:Accession: S57962  
R:Sanchez-Beato, A.; Garcia, J.  
submitted to the EMBL Data Library, July 1995

A:Description: Molecular characterization of a family of choline-binding proteins of *Cld*  
A:Reference number: S57714  
A:Accession: S57962  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-601 <S>  
A:Cross-references: EMBL:Z50033; NID:9899232; PIDN:CAA90347.1; PID:9899233  
C:Superfamily: cpl repeat homology  
F:533-552/Domain: cpl repeat homology <CP3>  
F:553-572/Domain: cpl repeat homology <CP4>

Query Match  
Best Local Similarity 27.0%; Score 83; DB 2; Length 601;  
Matches 31; Conservative 14; Mismatches 32; Indels 38; Gaps 6;

QY 57 NDEKSIQAEFLPMN-----NADTNQFATEKVTAVMGYNRE-----NAFRYET 102  
DB 324 DEQKNDKADFLNTGTGTINEDTKFVANGKLA-----YNTMNEKKVYVAYSLKT 378  
QY 103 EDGQVFTDVIAYSDNDV-----IYVPGTDGNEGYELWTTD-----YDN 143  
DB 379 KGGYYADEDESEKDEKDESSODDKTSVOTDVGN-----LWRLDGYIKKFDN 428

## RESULT 10

endo-1,4-beta-xylanase (EC 3.2.1.8) precursor, bifunctional - *Ruminococcus flavefaciens*  
N:Contains: endo-1,4-beta-xylanase (EC 3.2.1.8)  
C:Species: *Ruminococcus flavefaciens*  
C:Date: 22-Nov-1993 #sequence\_revision 22-Nov-1996 #text\_change 18-Jun-1999  
R:Zhang, J.X.; Flint, H.J.  
Mol. Microbiol. 6, 1013-1023, 1992  
A:Title: A bifunctional xylanase encoded by the xyna gene of the rumen cellulolytic bacterium.  
A:Reference number: S20907; MUID:92261318; PMID:1584021  
A:Accession: S20907  
A:Molecule type: DNA  
A:Residues: 1-954 <H>  
A:Cross-references: EMBL:Z11127; NID:946161; PIDN:CAA77476.1; PID:9581505  
A:Experimental source: strain 17  
A:Gene: xyna  
A:Genetics:  
A:Start codon: TTG  
A:Function:  
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic bonds in xylans  
A:Pathway: xylan degradation  
C:Superfamily: *Ruminococcus* bifunctional endo-1,4-beta-xylanase; endo-1,4-beta-xylanase  
C:Keywords: extracellular protein; glycosidase; hydrolase; multifunctional enzyme; polysaccharide  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-954/Product: bifunctional endo-1,4-beta-xylanase #status predicted <MAT>  
F:28-248/Domain: catalytic domain A #status predicted <CDA>  
F:40-236/Domain: endo-1,4-beta-xylanase homology <XYL>  
F:249-622/Domain: linker region B #status predicted <LRB>  
F:623-954/Domain: catalytic domain C #status predicted <CDC>  
F:655-953/Domain: streptomycin endo-1,4-beta-xylanase A homology <XYL>  
F:122,223/Active site: Glu #status predicted

Query Match  
Best Local Similarity 24.3%; Score 82.5; DB 1; Length 954;  
Matches 43; Conservative 26; Mismatches 37; Indels 71; Gaps 13;

QY 3 PDW---ADEANGAH---QDAW-KSLKADVENYVYKATYKNDPVWGNDFTCVGVANDV 56  
DB 722 PDWFFENSONGAYVSKDIMNQRELSMIKTFPAKSOYPNIDVSYD-VC----- 772  
QY 57 NDEKSIQAEFLPMNADTNQFATEKVTAVMGYNRE-----NAFRYETGQVFTDVIAY 114  
DB 773 NE-----LELNGG-GMRGA-DNSMVKIYGDSEFVIAFKYARQ-----Y 811  
QY 115 SDNCDVIYVPGTDGNEGYELWTTDYNIPANC-----LNKFNAYAVGRETRDVFSTA 168  
DB 812 A-----PACCKIYLDYNEY-IPAKTNDIYVNA 838

## RESULT 11

DNA-directed RNA polymerase (EC 2.7.7.6) I second largest chain - fruit fly (*Drosophila*)  
J00354  
C:Species: *Drosophila melanogaster*  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Jul-2000  
C:Accession: J00354; S07025  
R:Kontermann, R.; Stitzler, S.; Selfarth, W.; Petersen, G.; Bautz, E.K.F.  
Mol. Gen. Genet. 219, 373-380, 1989  
A:Title: Primary structure and functional aspects of the gene coding for the second-1  
A:Reference number: J00354; MUID:90158499; PMID:2482932  
A:Accession: J00354  
A:Molecule type: DNA  
A:Residues: 1-1129 <KON>  
A:Cross-references: EMBL:X17298; NID:98473; PIDN:CAA35185.1; PID:98474  
A:Note: the authors translated the codon CAG for residue 202 as Glu, TCT for residue  
C:Genetics:  
A:Gene: dmRP135  
A:Cross-references: FlyBase:FBgn0003278  
A:Introns: 30/3; 530/3  
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide  
C:Keywords: DNA binding; nucleotidyltransferase; nucleus; transcription; zinc finger

Query Match  
Best Local Similarity 25.0%; Score 81.5; DB 2; Length 1129;  
Matches 21; Conservative 15; Mismatches 37; Indels 11; Gaps 3;

QY 11 NGAHQDAWKSLK--DVENYVYKATYKNDPV-----W--GNDFTCVGVANDVNE 59  
DB 158 HGEHSEMGIRFYIRGNEKTVRLIMTRNHPICVRRSSWKDRQNSDLGLVQTVRED 217  
QY 60 EKSIQAEFLPMNADTNQFATEK 83  
DB 218 ESSLNSVHYHLNNGTAKMFMSHVK 241

## RESULT 12

myb-related transcription factor MYB14 [imported] - *Arabidopsis thaliana* (fragment)  
T51637  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 17-Nov-2000  
R:Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.  
Plant J. 16, 263-276, 1998  
A:Title: Towards functional characterisation of the members of the R2R3-MYB gene from  
A:Reference number: Z14349; MUID:9839469; PMID:9839469  
A:Accession: T51637  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-166 <KRA>  
A:Cross-references: EMBL:AF062865; PIDN:AAC3587.1  
A:Experimental source: cultivar Columbia  
A:Genetics:  
A:Gene: MYB14  
C:Superfamily: *Arabidopsis* 28K leaf-specific myb-related protein; myb DNA-binding rep  
C:Keywords: transcription factor

Query Match  
Best Local Similarity 22.6%; Score 80; DB 2; Length 166;  
Matches 38; Conservative 27; Mismatches 65; Indels 38; Gaps 7;

QY 18 WKSLK-----DVENYVYKATYKNDPVWGNDFTCVGVANDVNEDEKSIQAEFL 68  
DB 6 WSAIAKLPGRDNEIKNNWHTLKRSLKLNLSNGDPTADVNGINTTEDGSGVYVD-- 63  
QY 69 FMNADTNQFATEKVTAVMGYNRENAFRYETEDGQVFTDVIAYSDN--CDVITYPG 126  
DB 64 -----TASIQQFNSITFTDISNKKDINSYE-----DISALIDDSFMSDVYSVN 110  
QY 127 TDGNEGYELWTTDYNIPANC-----LNKFNAYAVGRETRDVFSTA 167

Db 111 SNRKEKIEDMEGLIDRNSKCSYNSKLYNDMEFNF-----DVFTS 153

## RESULT 13

probable MYB family transcription factor [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002

C:Accession: E84717

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Baito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Euse, D.; Nleman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402: 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84717

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-249 <STO>

A:Cross-references: GB:AE002093; MID:g4432813; PIDN:AAD20663.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g31180

A:Map position: 2

C:Superfamily: Arabidopsis 28k leaf-specific myb-related protein; myb DNA-binding repeat

Query Match 8.5%; Score 80; DB 2; Length 249;  
Best Local Similarity 22.6%; Pred. No. 8.4;  
Matches 38; Conservative 27; Mismatches 65; Indels 38; Gaps 7;

18 WKSLEK-----DVENVYVM-VKATYKNDPVMGNDFTCVGVANDVEDKSIQAEFL 68

Db 89 WSAIAAKLPRTDNEIRNVHTLKRSLKLNNGCJTKVNCINETNEDKSGSIYD-- 146

Qy 69 FMNADVDNMGAEKVTAVKMYGVNRENAFRYETEDGVFTVAYSDDN--CGVIYVPG 126

Db 147 -----TASLDQFSNITTFDISNDKNDIMSTF-----DISALIDDSWSDVISVN 193

Qy 127 TDGNEGEYELMTTDYDNIIPANC-----LNKFNVAVGRTRDVF 167

Db 194 SNRKEKIEDMEGLIDRNSKCSYNSKLYNDMEFNF-----DVFTS 236

RESULT 14  
A12159  
hypothetical protein alr2832 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C:Accession: A12159

R:Kaneke, T.; Nakamura, Y.; Wolik, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriyuchi, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8: 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: A12159

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-362 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA074531.1; PID:g17131926; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr2832

C:Superfamily: hypothetical protein alr1534

Query Match 8.5%; Score 79.5; DB 2; Length 362;

Best Local Similarity 23.6%; Pred. No. 15;

Matches 42; Conservative 26; Mismatches 69; Indels 41; Gaps 9;

Qy 23 ADVENVYVMKA---TYKNDPVMGNDFTCVG---VMANDVEDKSIQAEFL---FMNN 72

Db 192 SDTDEYTLASARMSITKRDIIYEAFLNWLGMPLISGDGPERRR-LQAKALDNKIFLGH 250

Qy 73 ADTNMQ--FATER--VTAVKMYGVNRENAFRYETEDGVFTVAYSDDNCDVIYVPG 126

Db 251 VSDNRKELEFSKAKSIIVALEYGL-----VPVEANASGTPVIAFGAGVLDITQING 303

Qy 127 TDG-----NEGEYELMTTDYDNIIPANC LNKNFENEVAVERTRDVF 169

Db 304 QTGVFKRQRPESLQKALLESGEITW--DYENIRNNAVNNSFEPFSKVERVITQIC 359

## RESULT 15

D90538

hypothetical protein MYPU\_2120 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

R:Chambud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, Nucleic Acids Res. 29: 2145-2153, 2001.

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p

A:Reference number: A9512; MUID:21267165; PMID:11353084

A:Accession: D90538

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1377 <KUR>

A:Cross-references: GB:AL445566; PID:g14089625; PIDN:CAC13385.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPU\_2120

A:Genetic code: SCC3

Query Match 8.5%; Score 79.5; DB 2; Length 1377;  
Best Local Similarity 24.7%; Pred. No. 83;  
Matches 42; Conservative 27; Mismatches 50; Indels 51; Gaps 11;

15 QDAWKSADVENYVMKAT-----YKND-----PYMGND-----FTCVGMAN-D 55

Db 1116 QDAWKS-----EDKVIYDOKANQNTIASQYKLDQKVKIRPNSDKALKEITFLGTTINTLT 1171

Qy 56 VNEDEKSIQAEFLFMNADTNMGDPATEKVTAVKMYGVNRENAFRYETEDGVFTVAYS 115

Db 1172 INED-KVWRADF-----AD-----DKRIDIKRYKLDLEIYSEFE-----GYT 1208

Qy 116 DDNCDVIYVFTDGNNEGEYELMTTDYDNIIPANC LNKNFENEVAVERTRDVF 165

Db 1209 NSNGEQLWVRNAQ-NEEAHAV---FNLLSSDNNSKYGYSLSVDQINAF 1253

Search completed: August 1, 2003, 12:16:34  
Job time : 14.7029 secs

**THIS PAGE BLANK (USPTO)**



```
RESULT 2
; Sequence 7, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 171
; TYPE: PRT
; ORGANISM: FS-HBP 2
US-10-087-195-7

Query Match
Best Local Similarity 100.0%; Score 936; DB 14; Length 171;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
Db 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
QY 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
Db 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 171
Db 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 171

RESULT 3
US-10-085-572-6
; Sequence 6, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGANISM: FS-HBP1
US-10-085-572-6

Query Match
Best Local Similarity 67.3%; Score 630; DB 14; Length 172;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
Db 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
QY 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
Db 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 171
Db 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 171

RESULT 4
US-10-087-195-6
; Sequence 6, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGANISM: FS-HBP1
US-10-087-195-6

Query Match
Best Local Similarity 67.3%; Score 630; DB 14; Length 172;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
Db 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
QY 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
Db 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 170
Db 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 170

RESULT 5
US-10-085-572-8
; Sequence 8, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 182
; TYPE: PRT
; ORGANISM: MS-HBP 1
US-10-085-572-8

Query Match
Best Local Similarity 35.4%; Score 331.5; DB 14; Length 182;
```

```
QY 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
Db 61 KNEVAMFEMFNADTYQHTEKATPDKMGYNKENATYQTEDEGVLTDLAFSDNCY 120
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 170
Db 121 VIYVPGTGDGAGYELMATDYTDVPASCLKFNENYAVGLPVRDVTSDCL 170

RESULT 4
US-10-087-195-6
; Sequence 6, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGANISM: FS-HBP1
US-10-087-195-6

Query Match
Best Local Similarity 67.3%; Score 630; DB 14; Length 172;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
Db 1 NOPDMADEAANGAHODAMKSLKADVENYVMVATKYNDPVGNDFTCVGMANDVNEDE 60
QY 61 KSIQAEFLFNNADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120
Db 61 KNEVAMFEMFNADTYQHTEKATPDKMGYNKENATYQTEDEGVLTDLAFSDNCY 120
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 170
Db 121 VIYVPGTGDGAGYELMATDYTDVPASCLKFNENYAVGLPVRDVTSDCL 170

RESULT 5
US-10-085-572-8
; Sequence 8, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 182
; TYPE: PRT
; ORGANISM: MS-HBP 1
US-10-085-572-8

Query Match
Best Local Similarity 35.4%; Score 331.5; DB 14; Length 182;
```



QY 127. TDGNESEYELMTDYDNIPANC-----LNKNEVAVGRETVDFTS 167  
Db 194 SNRNEKKEDEMEGLIDRNSKRCYSNSKLYNDMEFEV-----DVFTS 236

RESULT 9  
US-09-738-626-5144

; Sequence 5144, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIRO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 5144  
; LENGTH: 929  
; TYPE: PRN  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5144

Query Match 8.5%; Score 80; DB 10; Length 929;  
Best Local Similarity 25.6%; Pred. No. 7.4;  
Matches 31; Conservative 20; Mismatches 56; Indels 14; Gaps 3;

QY 21 LKADVENYVMATYATKNDPWCNDCTCVGMANDV--NDEKSIQAEFLFNNADTNMQ 78  
Db 73 IKALSNOKYHDVLAAGSD-----AVGLTGDSINHDADIVWTTTEVLNMYAGS 124  
QY 79 FATEKYAVKMYGYNRENAFRYETEDGOVFTDIYASDNCQVYIYPGTDGNEGEYELT 138  
Db 125 FALERLSHYV-----DEIHFLADSRGAVWEVYILNDSDVNTILGSAIVSNSEFEGL 180  
QY 139 T 139  
Db 181 T 181

RESULT 10  
US-09-815-242-5484

; Sequence 5484, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 1410  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5484  
; LENGTH: 327  
; TYPE: PRN  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-5484

Query Match 8.4%; Score 79; DB 9; Length 327;  
Best Local Similarity 23.1%; Pred. No. 2.4;  
Matches 39; Conservative 25; Mismatches 63; Indels 42; Gaps 8;

QY 14 HODAKSLKADVENYVMV---KATYKNDPWCNDCT--CVGMANDVNEDEKSIQAEFL 68  
Db 173 HKRAYRKEVEVEEYTYVPLGKADYKRE---GEDLTVPYCYGLAVNYCQADAILAADGI 229  
QY 69 FNNADTNMQFATEKYTAV-----KMGYNRENAFRYETEDGOVFTDIYASDNC-- 119  
Db 230 NVEYVDLRYTYPLDKETIIDRAKHGKVLVTFEDNL-----BGSINSEVSAILIARHCLF 283  
QY 120 -----DYIYVPGTDGNEGEYELMTTYDNIPANCCLKKFNEXA 156  
Db 284 ELDAPIRLAAPDVPSPMPSVLEN--EIMMN-----PEKILNKRRELA 325

## RESULT 11

US-09-815-242-12586  
; Sequence 12586, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 1410  
; SOFTWARE: FastSeq for Windows Version 4.0









STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,197C  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimdala, Michele A  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050, 041000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
US-08-282-197C-61

Query Match 8.8%; Score 82.5; DB 2; Length 405;  
Best Local Similarity 24.3%; Pred. No. 0.82;  
Matches 43; Conservative 26; Mismatches 37; Indels 71; Gaps 13;

QY 3 PDW---ADEAANGAH--QDAW-KSLKADVENYIMVATKNDPWCNGDTFCGVMANDV 56  
DB 172 PDWFRNFNSNGYYSKIDIMNORLESIMKTFALKSQYPNLDVYSYD-VC----- 222  
QY 57 NEDKSIQAEFLPMNADNMOPATEKYTAVKMYGVR--ENAFRYETEDQGVFTDYIAY 114  
DB 223 NE-----LPLANGG-GMRGA-DNSNWKYKIGDSDFYINAFKARQ-----Y 261  
QY 115 SDNCQVYVPGTGDNGEGYELMTTDYDNIIPANC--LNFENEYAVGRETRDVFYTS 168  
DB 262 A-----PACCKLYLNDYNEY-IPAKTNDIYNMA 288

RESULT 3  
PCT-US96-01314-56  
Sequence 56, Application PC/TUS9601314  
GENERAL INFORMATION:  
APPLICANT: M. Amin Arnaout  
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01314  
FILING DATE: 30-JAN-96  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/380,167  
FILING DATE: 30-JAN-95  
ATTORNEY/AGENT INFORMATION:  
NAME: John W. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/267001

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 252 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-01314-56

Query Match 8.4%; Score 78.5; DB 5; Length 252;  
Best Local Similarity 23.1%; Pred. No. 1.1;  
Matches 28; Conservative 20; Mismatches 50; Indels 23; Gaps 5;

QY 42 WGN-----DFTCGVANDYNEDEKSIQAEFLFN-----NADTNQFAT-----EKV 84  
DB 82 WPNSDPPEFSFKNVISTLEDVDEFRRNKLOGERISGNDAPBGGFDALIQAVCTRIDIGWRP 141  
QY 85 TAVKMYGVRNENAFRYETEDQGVFTDYIAYSDNCQVYVPGTGDNGEGYELMTTDYDNI 144  
DB 142 DSTHLVFTYESAFHYADGANVLAGTMSRDERCHL-----DTGTYYQYR--TQDIPSV 195  
QY 145 P 145  
DB 196 P 196

RESULT 4  
PCT-US93-11110-1  
Sequence 1, Application PC/TUS9311110  
GENERAL INFORMATION:  
APPLICANT: Schering Corp.  
TITLE OF INVENTION: Antagonists of Human Gamma Interferon  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering-Plough Corporation  
STREET: One Giralda Farms  
CITY: Madison  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07940  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 6.0.8  
SOFTWARE: Microsoft Word 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11110  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: U.S. Patent Application No. 07/980,527 filed November 2  
ATTORNEY/AGENT INFORMATION:  
NAME: Lunn, Paul G.  
REGISTRATION NUMBER: 32,743  
REFERENCE/DOCKET NUMBER: JB02850  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-822-7255  
TELEFAX: 201-822-7039  
TELEX: 219165  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 489 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-11110-1

Query Match 8.2%; Score 76.5; DB 5; Length 489;  
Best Local Similarity 19.3%; Pred. No. 5;  
Matches 37; Conservative 30; Mismatches 70; Indels 55; Gaps 11;

OY 4 DWAEANGAH-----ODAKSLKADY---ENYVYWKATYKNDPVGNDP- 46  
DB 72 EMIDACINISHHCNISDHVGDPSNSLWVRKARVGOKESAY-----AKSEFEA 120  
OY 47 TC-----VGVMANDVNEDEKSIOAEF---LFMNNADTMQOPATEKVTAVKMG-YNRENA 97  
DB 121 VCRGKIGPRLDIRKEKQIMIDIFHPSVFNVDGDEVDYDPETTCYIRVYNYVRANG 180  
OY 98 FRYETEDQVFTDVIAVSDNCVYIY---VPGTDGNE-----EG-YELMTTDDYDNI 146  
DB 181 -----SEIOYKILTOKEDEDCDEIQCOLAIPVSLNSQYCSAEGVLHVGVTTESKE 233  
OY 147 NCLKNFNEYAVG 158  
DB 234 VCITTFNSSING 245

RESULT 5  
US-08-871-572B-6  
Sequence 6, Application US/08871572B  
Patent No. 6287853  
GENERAL INFORMATION:  
APPLICANT: Pestka, Sidney  
APPLICANT: Kolenko, Serguei  
APPLICANT: Soh, Jaemog  
APPLICANT: Donnelly, Robert  
APPLICANT: Mariano, Thomas  
APPLICANT: Cook, Jeffrey  
APPLICANT: Emmanuel, Stuart  
APPLICANT: Schwartz, Barbara  
TITLE OF INVENTION: Accessory Factor for Interferon Gamma  
TITLE OF INVENTION: and its Receptor  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Richard R. Muccino  
STREET: 758 Springfield Avenue  
CITY: Summit  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07901  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/871,572B  
FILING DATE: 9-JUNE-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Muccino, Richard R.  
REGISTRATION NUMBER: 32,538  
REFERENCE/DOCKET NUMBER: UMD1-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 273-4988  
TELEFAX: (908) 273-4679  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
US-08-871-572B-6

Query Match 8.1%; Score 75.5; DB 3; Length 259;  
Best Local Similarity 19.3%; Pred. No. 2.5; Indels 55; Gaps 11;  
Matches 37; Conservative 29; Mismatches 71; Indels 55; Gaps 11;  
4 DWAEANGAH-----ODAKSLKADY---ENYVYWKATYKNDPVGNDP- 46

DB 72 EMIDACINISHHCNISDHVGDPSNSLWVRKARVGOKESAY-----AKSEFEA 120  
OY 47 TC-----VGVMANDVNEDEKSIOAEF---LFMNNADTMQOPATEKVTAVKMG-YNRENA 97  
DB 121 VCRGKIGPRLDIRKEKQIMIDIFHPSVFNVDGDEVDYDPETTCYIRVYNYVRANG 180  
OY 98 FRYETEDQVFTDVIAVSDNCVYIY---VPGTDGNE-----EG-YELMTTDDYDNI 146  
DB 181 -----SEIOYKILTOKEDEDCDEIQCOLAIPVSLNSQYCSAEGVLHVGVTTESKE 233  
OY 147 NCLKNFNEYAVG 158  
DB 234 VCITTFNSSING 245

RESULT 6  
US-08-190-204-2  
Sequence 2, Application US/08190204  
Patent No. 6558661  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Ward, Rebecca H.R.  
TITLE OF INVENTION: TREATMENT OF INFLAMMATORY BOWEL DISEASE WITH IFN-GAMMA  
TITLE OF INVENTION: INHIBITORS  
FILE REFERENCE: 11669, 90USMO  
CURRENT APPLICATION NUMBER: US/08/190,204  
CURRENT FILING DATE: 1994-02-22  
PRIOR APPLICATION NUMBER: 07/997,835  
PRIOR FILING DATE: 1992-12-29  
PRIOR APPLICATION NUMBER: PCT/US93/11966  
PRIOR FILING DATE: 1993-12-09  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 489  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-190-204-2

Query Match 8.1%; Score 75.5; DB 4; Length 489;  
Best Local Similarity 19.3%; Pred. No. 6.4;  
Matches 37; Conservative 29; Mismatches 71; Indels 55; Gaps 11;  
4 DWAEANGAH-----ODAKSLKADY---ENYVYWKATYKNDPVGNDP- 46  
DB 72 EMIDACINISHHCNISDHVGDPSNSLWVRKARVGOKESAY-----AKSEFEA 120  
OY 47 TC-----VGVMANDVNEDEKSIOAEF---LFMNNADTMQOPATEKVTAVKMG-YNRENA 97  
DB 121 VCRGKIGPRLDIRKEKQIMIDIFHPSVFNVDGDEVDYDPETTCYIRVYNYVRANG 180  
OY 98 FRYETEDQVFTDVIAVSDNCVYIY---VPGTDGNE-----EG-YELMTTDDYDNI 146  
DB 181 -----SEIOYKILTOKEDEDCDEIQCOLAIPVSLNSQYCSAEGVLHVGVTTESKE 233  
OY 147 NCLKNFNEYAVG 158  
DB 234 VCITTFNSSING 245

RESULT 7  
5221789-1  
Patent No. 5221789  
APPLICANT: NOVICK, DANIELA; RUBINSTEIN, MENACHEM  
TITLE OF INVENTION: INTERFERON-GAMMA RECEPTOR FRAGMENT AND  
ITS PRODUCTION  
NUMBER OF SEQUENCES: 1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/578,826  
FILING DATE: 07-SEP-1990  
SEQ ID NO: 1  
LENGTH: 489

5221789-1

Query Match	8.1%;	Score 75.5;	DB 6;	Length 489;
Best Local Similarity	19.3%;	Pred. No. 6.4;		
Matches 37;	Conservative 29;	Mismatches 71;	Indels 55;	Gaps 11.

```

QY      4 DAMDDEAANGH-----QDAKSLSKADY---ENYYMAKATYKNDPVMGDNF- 46
Db      72 EWDIDACINISHYKISDHVGPSPSLMVRKAFAGQKESAY-----AKSEFEA 120
QY      47 TC-----VGWANDVNEDEKSIQAEF-----LFNNADTQMCPATEKYTAVKMG- YNSENA 97
Db      121 VCDGKIGRPKIDIRKEEQIMIDLFHPSVFNPGEQGVYDDPPTCTCYIRYINYYAMNG 180
QY      98 FRFETEDGVFTDVIYASDNCQVYI---VPGTDGNF-----EG-YEIMTDDYDNIPIA 146
Db      181 -----SEIQRIILQKEDDCDEIQOLAIIPVSLNSQYCVSAEGVILHWGVTTESKE 233
QY      147 NCLNKFENEYAVG 158
Db      234 VCIITFNSSITG 245

```

RESULT 8  
US-08-79

Sequence 3, Application US/08790912  
Patent No. 5976542

GENERAL INFORMATION:  
APPLICANT: Weiser, Jeffrey N.  
APPLICANT: Plaut, Andrew G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT  
TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PAULITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
City: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-4U1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-464

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1964 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
IS-08-790-912-3

Query Match	8.0%;	Score 74.5;	DB 2;	Length 1964;
Best Local Similarity	25.2%;	Pred. No. 63;		
Matches	27;	Conservative	19;	Mismatches 32;
			Indels	29;
			Gaps	5;

```

Oy  61 KSIAEEFFMNADTNMQFAFEKTYAVGCGINRNARREYETE-----DGOVFPVI 112
Db  1124 KNIETLLEFYN-----KSTIVATGVLVENSLLYKELLSSAYMMKDDGYIIDIV 1172
Oy  113 AYSDNDGVIVYPGDCNBEGLT-WTIDYDNIIPANCINKRNEKVAVG 158
Db  1173 S-NQOTAKLHLHYNDHSSEKFDLQYQIDFNP-----ETNLG 1210

```

## RESULT 9

```

US-08-790-912-2
Sequence 2, Application US/08790912
Patent No. 5976542
GENERAL INFORMATION:
Applicant: Welser, Jeffrey N.
Applicant: Plaut, Andrew G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/790,912
FILING DATE: 29-JAN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,939
FILING DATE: 23-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: leary, Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 7600-401
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2052 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-790-912-2

```

Query Match	8.0%;	Score 74.5;	DB 2;	Length 2052;
Best Local Similarity	25.2%;	Pred. No. 67;		
Matches 27;	Conservative 19;	Mismatches 32;	Indels 29;	Gaps 5

```

OY 61 KSIQAELENNADTNMQFAFEKTAIVAMCYGIRNNARREYEE-----DGQFEYVI 112
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 1197 KNIRKLKLPFY-----KSTIYKGYGVAKVENSLLYQKELLASVMMKDDQYITDIY 1244
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

OY 113 AYSDNDQVIYVPGTDCGNEGIEL-WTDDYDNIIPANCLINKENBAVG 158
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 1246 S-NKQTANKLLLHYNDHSSEKFDIKYQDTEANLP-----EYVLG 1283
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

## RESULT 10

US-09-107-532A-5946  
Sequence 5946, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5946:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: (B) LOCATION 1...420  
SEQUENCE DESCRIPTION: SEQ ID NO: 5946:  
US-09-107-532A-5946  
Query Match 7.9%; Score 74; DB 4; Length 420;  
Best Local Similarity 23.3%; Pred. No. 7.5;  
Matches 41; Conservative 15; Mismatches 58; Indels 62; Gaps 9;  
QY 4 DWADGANGAHQDAWKSIAKADVENYVYVAVKATYKNDPYWGNDFTCVGMANDV-----N 57  
DB 261 DWDD-----MLDVATNYPAPAFETIKQDQDQGE-----GQMIDDIYQILGK 301  
QY 58 EDEKSIQAEF---LFMNNADTNMQFATEKTAIVAMGYNKNMARFETEDGQVTTDVYAY 114  
DB 302 NDLLSTONELKKQINIKTSITLPLNKN-----NASMFEFDSDD----- 342  
QY 115 SDNCDVYVPGTGNEGEYELMTTDYDNIIPANCLNKNE-----YAVGRETRDVF 166  
DB 343 --DSTN-----GDGSNSBEGY---TDYSSISENTYDQIDPNYTYQOTEDTSQYTT 387

RESULT 11  
US-09-107-532A-6865  
Sequence 6865, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 6865:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 113 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: (B) LOCATION 1...113  
SEQUENCE DESCRIPTION: SEQ ID NO: 6865:  
US-09-107-532A-6865  
Query Match 7.9%; Score 73.5; DB 4; Length 113;  
Best Local Similarity 29.4%; Pred. No. 1.3;  
Matches 32; Conservative 8; Mismatches 54; Indels 15; Gaps 5;  
QY 67 FLFNNADTNMQFATEKTAIVAMGYNKNMARFETEDGQVTTDVYAVS-----DDNCDV 121  
DB 1 FQVHEVADTMRLRYRPFKNKDSGYELSKTIRERTINDSYIKETILYSVARLEKNOLI 60  
QY 122 IYVPG--TDGNEGEYELMTTDYDNIIPANCL-NKNEYAVGRETRDVF 167  
DB 61 VSYPGEVTHGKRRTYRIT---EIGRNLLEKKQEWLT---TKDVVTS 102

RESULT 12  
US-09-107-532A-4711  
Sequence 4711, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <unknown>

SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107.532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariadello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 4711:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 489 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (B) LOCATION 1...489  
SEQUENCE DESCRIPTION: SEQ ID NO: 4711:  
US-09-107-532A-4711

Query Match 7.9%; Score 73.5; DB 4; Length 489;  
Best Local Similarity 23.5%; Pred. No. 11;  
Matches 32; Conservative 19; Mismatches 30; Indels 55; Gaps 8;  
Db 44 DFEETPIVETATSDLRPSIDPV-----ILI-----KLVIQYLFGRSMRQT 87  
QY 24 DVENVYIMKATYKN-----DPWGNDFTCVGMANDVDEKSIQAEFLF----- 69  
Db 70 MNNDVPMQPA-----TEKVTAVKMYGYNRENAFRETEDGQFTDVAY----- 114  
QY 88 IKEDFTVAARWFLGYSFEKIKHFSFGKMYRRPRETT-----VFEDITAYILQAVKA 143  
Db 115 ---SDNCQDVIVPGT 127  
QY 144 GFVTEEDN---LYLDST 156  
Db

RESULT 13  
US-09-134-001C-3716  
Sequence 3716, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3716  
LENGTH: 1335  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3716

Query Match 7.8%; Score 73; DB 4; Length 1335;  
Best Local Similarity 23.7%; Pred. No. 53;  
Matches 36; Conservative 18; Mismatches 46; Indels 52; Gaps 9;

QY 19 KSLKAD-VENY--YMKATYKND---PWGNDFTCVGMANDVNEDEKSIQAEFLPMNN 72  
Db 329 KMKAPRIEEDYTSYFPRYGRNGVRP-----EGIVYHDPRANDNSTIIDGLAFMKR 380  
QY 73 ADTNMOPATEKVTAVKMYGYNRENAFRETEDGQV-----TDVIASDNCQDVIVPGT 127  
Db 381 -----NYTNAFVHARVDGNRIETAPTDYLSWG-----AGP 411  
QY 128 DGNEGYE---LWTTDYDNIPANCLNFNEYA 156  
Db 412 YGNORFIVEIVHVDHDSF-ARSMNNYADYA 442  
RESULT 14  
5268463-2  
Patent No. 5268463  
APPLICANT: JEFFERSON, RICHARD A.  
TITLE OF INVENTION: PLANT PROMOTER a-GLUCURONIDASE GENE  
CONSTRUCT  
NUMBER OF SEQUENCES: 9  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/447,976  
FILING DATE: 08-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 119,102  
FILING DATE: 10-NOV-1987  
APPLICATION NUMBER: 264,586  
FILING DATE: 31-OCT-1988  
SEQ ID NO: 2  
LENGTH: 600  
5268463-2

Query Match 7.7%; Score 72.5; DB 6; Length 600;  
Best Local Similarity 23.4%; Pred. No. 19;  
Matches 41; Conservative 26; Mismatches 55; Indels 53; Gaps 10;  
Db 1 NOPD-WADERANG---AHODAMKSLKA-----DVE-----NYVYV 32  
QY 368 NKRELYSEAVANGETQOAHLOAKELIANDKNHPSVMMSTANEDDTROVHGNSPLA 427  
Db 33 KATYKNDPWGNDFTCVGMANDVNEDEKSIQAEFLPMNANDTNMOPATEKVTAVKMYG 92  
QY 428 EATRKLDPT--RPTCVNWFCDARHDTISDLDFVLCN-----RYGCV 469  
Db 93 NRENAFRETEDGQFTDVAYASDNCQDVIVP--GTD---GNEBEG-ELMTDY 141  
QY 470 YVDSG-DLETAERVLKELIAWOKLHOPITITTEYGVDTLAGHSMYTDWMSSEY 523  
Db

RESULT 15  
US-08-882-704A-5  
Sequence 5, Application US/08882704A  
Patent No. 5879906  
GENERAL INFORMATION:  
APPLICANT: Jefferson, Richard A.  
APPLICANT: Wilson, Katherine J.  
APPLICANT: Leader, Michael  
TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/08/882,704A  
 FILING DATE: 25-JUN-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 587906tendburg Ph.D., Carol  
 REGISTRATION NUMBER: 39,317  
 REFERENCE/DOCKET NUMBER: 190106.404  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 602 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-882-704A-5

Query Match 7.7%; Score 72.5; DB 2; Length 602;  
 Best Local Similarity 23.4%; Pred. No. 19;  
 Matches 41; Conservative 26; Mismatches 55; Indels 53; Gaps 10;

OY 1 NQPD--WADEANG---AHODAMKSLKA-----DVE-----NYYMV 32  
 Db 369 NKPKELYSEAVNGETOQAHLOAIKELIARDKNHPSVYVMSIANEPDTRPQVHGNIISPLA 428  
 OY 33 KATYKNDPVNGNDFTCVGANVDYNEDEKSTQAEFLFMNNADTNMGEATEKVTAVKMYGY 92  
 Db 429 EATRKLDPT--RPTTCVNMVFCDAHTDTISDLFDVLCIN-----RYTGW 470  
 OY 93 NRENAPRYETEDCGVFTDVIAYSDDNCDVIYVP--GTD--GNEEGY-ELWTTDY 141  
 Db 471 YVQSG-DLEFTEKYLEKELAMQEKLNQPIITTEYGVDTLAGLHSMYTDMSSEY 524

Search completed: August 1, 2003, 12:17:20  
 Job time : 14.0286 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 : Search time 53.0914 Seconds

(Without alignments)  
511.236 Million cell updates/sec

Title: US-10-087-195-7

Perfect score: 936  
Sequence: 1 NOPDMADEAANGAHQDAWKS.....FNEYAVGRETRDYFTSACLE 171

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A\_Geneseq\_19Jun03.\*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	100.0	171	22	AA1980.DAT.*
2	936	100.0	171	22	AA1981.DAT.*
3	936	100.0	190	19	AA1982.DAT.*
4	936	100.0	190	20	AA1983.DAT.*
5	630	67.3	172	22	AA1984.DAT.*
6	630	67.3	172	22	AA1985.DAT.*
7	630	67.3	190	19	AA1986.DAT.*
8	630	67.3	190	20	AA1987.DAT.*
9	331.5	35.4	182	22	AA1988.DAT.*

10	331.5	35.4	182	22	AA1989.DAT.*
11	331.5	35.4	200	19	AA1990.DAT.*
12	331.5	35.4	200	20	AA1991.DAT.*
13	288	30.8	203	20	AA1992.DAT.*
14	285	30.4	203	20	AA1993.DAT.*
15	282	30.1	209	19	AA1994.DAT.*
16	282	30.1	209	20	AA1995.DAT.*
17	222	23.7	198	20	AA1996.DAT.*
18	113	12.1	285	20	AA1997.DAT.*
19	105	11.2	1592	14	AA1998.DAT.*
20	99	10.6	284	20	AA1999.DAT.*
21	91.5	9.8	221	22	AA2000.DAT.*
22	87	9.3	338	22	AA2001.DAT.*
23	86.5	9.2	738	23	AA2002.DAT.*
24	85.5	9.1	337	23	AA2003.DAT.*
25	85.5	9.1	1612	23	AA2004.DAT.*
26	84.5	9.0	408	22	AA2005.DAT.*
27	84	9.0	624	17	AA2006.DAT.*
28	83.5	8.9	249	21	AA2007.DAT.*
29	83.5	8.9	275	21	AA2008.DAT.*
30	83.5	8.9	2911	24	AA2009.DAT.*
31	83.5	8.9	2912	22	AA2010.DAT.*
32	82.5	8.8	405	17	AA2011.DAT.*
33	82.5	8.8	954	21	AA2012.DAT.*
34	82.5	8.8	954	23	AA2013.DAT.*
35	81.5	8.7	1129	22	AA2014.DAT.*
36	80.5	8.6	516	22	AA2015.DAT.*
37	80	8.5	240	21	AA2016.DAT.*
38	80	8.5	249	21	AA2017.DAT.*
39	80	8.5	249	22	AA2018.DAT.*
40	80	8.5	321	20	AA2019.DAT.*
41	80	8.5	929	22	AA2020.DAT.*
42	79	8.4	327	22	AA2021.DAT.*
43	79	8.4	327	22	AA2022.DAT.*
44	78.5	8.4	1752	24	AA2023.DAT.*
45	78.5	8.4	1752	24	AA2024.DAT.*

#### ALIGNMENTS

RESULT 1  
AAB74289 standard; protein: 171 AA.  
AAB74289;  
20-JUN-2001 (first entry)  
Histacalain protein FS-HBP2.  
Histacalain; FS-HBP1; conjunctivitis.  
Unidentified.  
W0200115719-A2.  
08-MAR-2001.  
24-AUG-2000; 2000WO-GB03282.  
01-SEP-1999; 99GB-0020674.  
(EVOL-) EVOLUTEC LTD.  
Nutall PA, Paesen GC;  
WPI, 2001-257675/26.  
Use of histacalain proteins for treating or preventing non-infective conjunctivitis, or for manufacturing a medicament for treating or preventing conjunctivitis, e.g. seasonal or perennial allergic conjunctivitis

XX Claim 4; Page 5-6; 19pp; English.  
PS  
CC The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein-FS-HBP1. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.  
XX  
SQ Sequence 171 AA;  
Query Match 100.0%; Score 936; DB 22; Length 171;  
Best Local Similarity 100.0%; Pred. No. 1.6e-89;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NOPDMADEAANGAHODAMKSLKADVENYVYKATYKNDPVWGNDFTCVGMANDVNEDE 60  
DB 1 NOPDMADEAANGAHODAMKSLKADVENYVYKATYKNDPVWGNDFTCVGMANDVNEDE 60  
QY 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
DB 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
QY 121 VIIVPGTGNBEGYEELMTTDDYDNIIPANCLKFNFEYAVGRTDVFYSACLE 171  
DB 121 VIIVPGTGNBEGYEELMTTDDYDNIIPANCLKFNFEYAVGRTDVFYSACLE 171  
RESULT 2  
AAB73261  
ID AAB73261 standard; protein; 171 AA.  
XX  
AC AAB73261;  
XX  
DT 15-MAY-2001 (first entry)  
XX  
DE Histacalin protein FS-HBP2.  
XX  
KW FS-HBP2; histacalin protein; antinflammatory; antiallergic;  
KW ophthalmological; allergic rhinitis.  
XX  
OS Unidentified.  
XX  
PN WO200116164-A2.  
XX  
PD 08-MAR-2001.  
XX  
PF 24-AUG-2000; 2000WO-GB03287.  
XX  
PR 01-SEP-1999; 99GB-0020673.  
XX  
PA (EVOU- ) EVOLUTEC LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI; 2001-218521/22.  
XX  
PT Use of histacalin proteins for treating or preventing allergic  
PT rhinitis, or for manufacturing a medicament for treating or preventing  
PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis  
XX  
PS Disclosure; Pages 4-6; 19pp; English.  
XX  
CC The present invention relates to a method for treating or preventing  
CC allergic rhinitis. The method involves employing a blood-feeding  
CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
CC sequence is one such histacalin protein. The histacalin protein, is  
CC useful for treating or preventing allergic rhinitis, both seasonal and  
CC perennial allergic conjunctivitis.  
XX  
SQ Sequence 171 AA;  
Query Match 100.0%; Score 936; DB 22; Length 171;  
Best Local Similarity 100.0%; Pred. No. 1.6e-89;

Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NOPDMADEAANGAHODAMKSLKADVENYVYKATYKNDPVWGNDFTCVGMANDVNEDE 60  
DB 1 NOPDMADEAANGAHODAMKSLKADVENYVYKATYKNDPVWGNDFTCVGMANDVNEDE 60  
QY 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
DB 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
QY 121 VIIVPGTGNBEGYEELMTTDDYDNIIPANCLKFNFEYAVGRTDVFYSACLE 171  
DB 121 VIIVPGTGNBEGYEELMTTDDYDNIIPANCLKFNFEYAVGRTDVFYSACLE 171  
RESULT 3  
AAW37447  
ID AAW37447 standard; protein; 190 AA.  
XX  
AC AAW37447;  
XX  
DT 08-JUN-1998 (first entry)  
XX  
DE Tick vasoactive amine binding protein 2 FS-HBP2.  
XX  
KW Female-specific vasoactive amine binding protein 1; FS-HCP1;  
KW histamine; serotonin; assay; antihistamine; anti-inflammatory;  
KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
KW transgenic animal; tick.  
XX  
OS Rhipicephalus appendiculatus.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
FT /label= Sig\_peptide  
XX  
PN WO9744451-A2.  
XX  
PD 27-NOV-1997.  
XX  
PF 19-MAY-1997; 97WO-GB01372.  
XX  
PR 18-APR-1997; 97GB-0007844.  
XX  
PR 18-MAY-1996; 96GB-0010484.  
XX  
PA (OXFO- ) OXFORD VACS LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI; 1998-018506/02.  
XX  
DR N-PSDB; AAV00228.  
XX  
PT New vasoactive amine binding proteins and related nucleic acid,  
PT vectors - transformed cells and transgenic animals, used for  
PT assaying or removing histamine and as antihistamine or  
PT anti-inflammatory agents  
XX  
PS Example 2; Fig 2; 44pp; English.  
XX  
CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
CC female-specific histamine binding protein 2 (FS-HBP2). Its amino  
CC acid sequence was deduced from a cDNA clone (see AAV00228) obtained  
CC from a salivary gland cDNA library. FS-HBP1 (see AAW37446) and  
CC male-specific HSP1 (see AAW37448) and a related protein, D-RET6 (see  
CC AAW37449) from Dermacentor reticulatus, were also identified. These  
CC novel vasoactive amine binding proteins (VABPs) can be expressed  
CC in host cells using e.g. a baculovirus expression system. They can  
CC be used: (i) to assay histamine (or other VA such as serotonin) in  
CC body fluids or cell culture supernatants, e.g. to monitor the  
CC effect of allergens; (ii) for binding VA, e.g. to remove histamine  
CC from blood, food, cell cultures etc.; (iii) as an antihistamine or  
CC anti-inflammatory agents, e.g. for treating insect, snake or  
CC scorpion bites or dermatitis, or as a carrier for slow release of

CC histamine-related compounds; (iv) in vaccines to protect against  
CC metazoan parasites, especially in animals; (v) as reagents for  
CC studying inflammation, involvement of VA in ulcer formation or the  
CC immune response etc. VABPs provide a more sensitive assay for  
CC histamine than low-affinity antibodies currently used. They may  
CC also be more effective and safer than conventional antihistamines.

XX Sequence 190 AA;

Query Match 100.0%; Score 936; DB 19; Length 190;  
Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NOPMDADANCAHODAKSLKADVENYVYWKATYKKNDPVGNDFTCVGVANDVNEDE 60  
DB 20 NOPMDADANCAHODAKSLKADVENYVYWKATYKKNDPVGNDFTCVGVANDVNEDE 79  
OY 61 KSIQAEFLFMNNADTNMQFATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
DB 80 KSIQAEFLFMNNADTNMQFATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 139  
OY 121 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 171  
DB 140 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 190

#### RESULT 4

AAV18079  
ID AAV18079 standard; Protein; 190 AA.

AC AAV18079;

DT 06-AUG-1999 (first entry)

DE Histamine binding protein FS-HBP2.

XX Histamine binding protein; serotonin binding compound; inflammation;

KW gastric acid secretion; allergy; type I hypersensitivity reaction;

KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;

KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;

KW respiratory disease; coronary heart disease; cellular growth regulator;

KW tissue repair; blood-sucking ectoparasite; therapy.

XX Rhipicephalus appendiculatus.

OS W09927104-A1.

PN 03-JUN-1999.

PD 26-NOV-1998; 98MO-GB03530.

PF 26-JUN-1998; 98GB-0013917.

PR 26-NOV-1997; 97GB-0025046.

XX (OXFO-) OXFORD VACS LTD.

PA Nuttall PA, Paesen GC;

PI WPI: 1999-357641/30.

XX N-PSDB: AAV76965.

DR Histamine and serotonin binding compounds useful for the treatment

XX of allergies

PS Claim 1; Fig 2; 84pp; English.

CC This sequence is an example of a histamine or serotonin binding

CC compound (A), of the invention. The compounds are useful for regulating

CC the action of histamine and serotonin (in e.g. inflammation and gastric

CC acid secretion), the detection, quantification and removal of histamine

CC or serotonin (in animals, plants, cell cultures, food materials, or

CC humans) and in the treatment of various diseases and allergies

CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic

CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
CC allergies, abnormal blood pressure, migraine, psychological disorders,  
CC respiratory disease, and coronary heart disease). Histamine may also be  
CC used to regulate cellular growth and tissue repair. The molecules may  
CC also be used as components of vaccines directed against blood-sucking  
CC ectoparasites.

XX Sequence 190 AA;

Query Match 100.0%; Score 936; DB 20; Length 190;  
Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NOPMDADANCAHODAKSLKADVENYVYWKATYKKNDPVGNDFTCVGVANDVNEDE 60  
DB 20 NOPMDADANCAHODAKSLKADVENYVYWKATYKKNDPVGNDFTCVGVANDVNEDE 79  
OY 61 KSIQAEFLFMNNADTNMQFATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
DB 80 KSIQAEFLFMNNADTNMQFATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 139  
OY 121 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 171  
DB 140 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 190

#### RESULT 5

AAB74288  
ID AAB74288 standard; Protein; 172 AA.

AC AAB74288;

DT 20-JUN-2001 (first entry)

DE Histacalin protein FS-HBPL.

XX Histacalin; FS-HBPL; conjunctivitis.

OS Unidentified.

XX W0200115719-A2.

PN 08-MAR-2001.

PD 24-AUG-2000; 2000MO-GB03282.

PF 01-SEP-1999; 99GB-0020674.

PR (EVOL-) EVOLUTEC LTD.

PA Nuttall PA, Paesen GC;

PI WPI: 2001-257675/26.

XX Use of histacalin proteins for treating or preventing non-infective

XX conjunctivitis, or for manufacturing a medicament for treating or

XX preventing conjunctivitis, e.g. seasonal or perennial allergic

XX conjunctivitis

XX Claim 4; page 5-6; 19pp; English.

XX The present invention relates to the use of a histacalin protein

XX for treating or preventing conjunctivitis. The present sequence

XX is the histacalin protein FS-HBPL. The invention is particularly useful

XX in the treatment of allergic or seasonal conjunctivitis.

XX Sequence 172 AA;

Query Match 67.3%; Score 630; DB 22; Length 172;

Best Local Similarity 66.5%; Pred. No. 1.3e-57;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

OY 1 NOPMDADANCAHODAKSLKADVENYVYWKATYKKNDPVGNDFTCVGVANDVNEDE 60





XX AAB73262;  
 AC 15-MAY-2001 (first entry)  
 DT XX  
 DE Histacalin protein MS-HBPI.  
 KM MS-HBPI; histacalin protein; antiinflammatory; antiallergic;  
 KW ophthalmological; allergic rhinitis.  
 OS Unidentified.  
 PN W0200116164-A2.  
 PD 08-MAR-2001.  
 PF 24-AUG-2000; 2000WO-GB03287.  
 PR 01-SEP-1999; 99GB-0020673.  
 XX (EVOL-) EVOLUTEC LTD.  
 PA Nuttall PA, Paesen GC;  
 PI WPI; 2001-218521/22.  
 DR Use of histacalin proteins for treating or preventing allergic  
 XX rhinitis, or for manufacturing a medicament for treating or preventing  
 PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis -  
 PS Disclosure; Pages 4-6; 19pp; English.  
 XX The present invention relates to a method for treating or preventing  
 CC allergic rhinitis. The method involves employing a blood-feeding  
 CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
 CC sequence is one such histacalin protein. The histacalin protein, is  
 CC useful for treating or preventing allergic rhinitis, both seasonal and  
 CC perennial allergic conjunctivitis.  
 XX SQ Sequence 182 AA;  
 Query Match 35.4%; Score 331.5; DB 22; Length 182;  
 Best Local Similarity 42.0%; Pred. No. 2e-26;  
 Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
 QY 3 PMWADANAGAHODAMKSLKADVENYYVWKATYKNDPVGNDFTCVGVANDVNDKES 62  
 DB 2 PTWANEAKLGSYODAKSLQODONKRYTLAQATOTDVGVEFTCVSYAEKIG--KKK 59  
 QY 63 IOAEFLPMNADTNMOPATEKTVAVKMGYNRENAREYETED--GGVFTDVIAYSD-DNC 119  
 DB 60 LNAFTILYKKNHLLDLKESHETITVWKAYDYTTENGKRYETGRTGFEDVFYFSDYKNC 119  
 QY 120 DVIYVPGTDGNEG-YELWTTD--YDNIPANCLNKREYAVGNE--TRDVFY-SAC 169  
 DB 120 DVIYVPERKSGDEGDELWVSEDKIDPCCKFTMAYFAQOQEKTVRNVTYDSSC 175  
 RESULT 11  
 AAM37448 standard; Protein; 200 AA.  
 AC AAM37448;  
 DT 08-JUN-1998 (first entry)  
 DE Tick vasoactive amine binding protein 1 MS-HBPI.  
 XX Male-specific vasoactive amine binding protein 1; MS-HBPI;  
 KM histamine; serotonin; assay; antihistamine; anti-inflammatory;  
 KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
 XX transgenic animal; tick.

OS Rhipicephalus appendiculatus.  
 FH Key Location/Qualifiers  
 FT Peptide 1..18  
 FT Modified-site /label= Sig-peptide  
 FT 79..81  
 FT /note= "Asn is N-glycosylated"  
 PN W09744451-A2.  
 PD 27-NOV-1997.  
 PF 19-MAY-1997; 97WO-GB01372.  
 PR 18-APR-1997; 97GB-0007844.  
 PR 18-MAY-1996; 96GB-0010484.  
 XX (OXFO-) OXFORD VACS LTD.  
 PA Nuttall PA, Paesen GC;  
 PI WPI; 1998-018506/02.  
 DR N-PSDB; AAV00229.  
 XX New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 XX Example 2; Fig 3; 44pp; English.  
 PS This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
 CC male-specific histamine binding protein 1 (MS-HBPI). Its amino  
 CC acid sequence was deduced from a cDNA clone (see AAV00227) obtained  
 CC from a salivary gland cDNA library. Female-specific HSP1 and  
 CC HSP2 (see AAM37446-47) and a related protein, D.RET6 (see AAM37449)  
 CC from Dermacentor reticulatus, were also identified. These novel  
 CC vasoactive amine binding proteins (VABPs) can be expressed in  
 CC host cells using e.g. a baculovirus expression system. They can  
 CC be used: (i) to assay histamine (or other VA such as serotonin) in  
 CC body fluids or cell culture supernatants, e.g. to monitor the  
 CC effect of allergens; (ii) for binding VA, e.g. to remove histamine  
 CC from blood, food, cell cultures etc.; (iii) as an antihistamine or  
 CC anti-inflammatory agents, e.g. for treating insect, snake or  
 CC scorpion bites or dermatitis; or as a carrier for slow release of  
 CC histamine-related compounds; (iv) in vaccines to protect against  
 CC metazoan parasites, especially in animals; (v) as reagents for  
 CC studying inflammation. Involvement of VA in ulcer formation or the  
 CC immune response etc. VABPs provide a more sensitive assay for  
 CC histamine than low-affinity antibodies currently used. They may  
 CC also be more effective and safer than conventional antihistamines.  
 XX SQ Sequence 200 AA;  
 Query Match 35.4%; Score 331.5; DB 19; Length 200;  
 Best Local Similarity 42.0%; Pred. No. 2.3e-26;  
 Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
 QY 3 PMWADANAGAHODAMKSLKADVENYYVWKATYKNDPVGNDFTCVGVANDVNDKES 62  
 DB 20 PTWANEAKLGSYODAKSLQODONKRYTLAQATOTDVGVEFTCVSYAEKIG--KKK 77  
 QY 63 IOAEFLPMNADTNMOPATEKTVAVKMGYNRENAREYETED--GGVFTDVIAYSD-DNC 119  
 DB 78 LNAFTILYKKNHLLDLKESHETITVWKAYDYTTENGKRYETGRTGFEDVFYFSDYKNC 137  
 QY 120 DVIYVPGTDGNEG-YELWTTD--YDNIPANCLNKREYAVGNE--TRDVFY-SAC 169  
 DB 138 DVIYVPERKSGDEGDELWVSEDKIDPCCKFTMAYFAQOQEKTVRNVTYDSSC 193  
 RESULT 12  
 AAY18080



ID AAY18080 standard; Protein: 200 AA.  
XX  
AC AAY18080;  
XX  
DT 06-AUG-1999 (first entry)  
XX  
DE Histamine binding protein MS-HBPL.  
XX  
KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.  
XX  
OS Bhocephalus appendiculatus.  
XX  
PN MO9927104-A1.  
XX  
PD 03-JUN-1999.  
XX  
PE 26-NOV-1998; 98WO-GB03530.  
XX  
PF 26-JUN-1998; 98GB-0013917.  
XX  
PR 26-NOV-1997; 97GB-0025046.  
XX  
PA (OXFO-) OXFORD VACS LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI: 1999-357841/30.  
XX  
N-PSDB: AAX76966.  
XX  
PT Histamine and serotonin binding compounds useful for the treatment  
XX  
PT of allergies  
XX  
PS Claim 1; Fig 3; 84pp; English.  
XX  
XX This sequence is an example of a histamine or serotonin binding  
XX compound (A), of the invention. The compounds are useful for regulating  
XX the action of histamine and serotonin (in e.g. inflammation and gastric  
XX acid secretion), the detection, quantification and removal of histamine  
XX or serotonin (in animals, plants, cell cultures, food materials, or  
XX humans) and in the treatment of various diseases and allergies  
XX (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
XX rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
XX allergies, abnormal blood pressure, migraine, psychological disorders,  
XX respiratory disease, and coronary heart disease). Histamine may also be  
XX used to regulate cellular growth and tissue repair. The molecules may  
XX also be used as components of vaccines directed against blood-sucking  
XX ectoparasites.  
XX  
SQ Sequence 200 AA;  
XX  
Query Match 35.4%; Score 331.5; DB 20; Length 200;  
Best Local Similarity 42.0%; Pred. No. 2.3e-26;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
OY 3 PDMADEANGAHODAMKSLKADVENYVYKATYKNDPVGNDFTGCVGMANDVNEDEKS 62  
DB 20 PTMADEANGAHODAMKSLKADVENYVYKATYKNDPVGNDFTGCVGMANDVNEDEKS 77  
OY 63 IOAEFLPMNADTMNOFATEKVTAVKMGYRENAFRETEDE--GQVTDVIANSD-DNC 119  
DB 78 LNAITLYKNKHLTDLKESHERITWKADYTTENGIKETGTROTDEDVFESDYDNC 137  
OY 120 DVIVPPTDGNBEG-YELMTD--YDNIPANCLKNFENYAVGRE--TRDVPF-SAC 169  
DB 138 DVIVPPTDGNBEG-YELMTD--YDNIPANCLKNFENYAVGRE--TRDVPF-SAC 193  
RESULT 13  
AAY18085

ID AAY18085 standard; Protein: 203 AA.  
XX  
AC AAY18085;  
XX  
DT 06-AUG-1999 (first entry)  
XX  
DE Histamine binding protein 1h/Bm-HBPL.  
XX  
KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.  
XX  
OS Bophthalmus microplus.  
XX  
OS Ixodes hexagonus.  
XX  
OS Synthetic.  
XX  
PN MO9927104-A1.  
XX  
PD 03-JUN-1999.  
XX  
PE 26-NOV-1998; 98WO-GB03530.  
XX  
PF 26-JUN-1998; 98GB-0013917.  
XX  
PR 26-NOV-1997; 97GB-0025046.  
XX  
PA (OXFO-) OXFORD VACS LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI: 1999-357841/30.  
XX  
N-PSDB: AAX76970.  
XX  
PT Histamine and serotonin binding compounds useful for the treatment  
XX  
PT of allergies  
XX  
PS Claim 13; Fig 7; 84pp; English.  
XX  
XX This sequence is an example of a histamine or serotonin binding  
XX compound (A), of the invention. cDNA encoding this sequence was isolated  
XX from a mixed Bophthalmus microplus/Ixodes hexagonus cDNA expression  
XX library. The compounds are useful for regulating the action of histamine  
XX and serotonin (in e.g. inflammation and gastric acid secretion), the  
XX detection, quantification and removal of histamine or serotonin (in  
XX animals, plants, cell cultures, food materials, or humans) and in the  
XX treatment of various diseases and allergies (e.g. type I hypersensitivity  
XX reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic  
XX dermatitis, insect bites and food and drug allergies, abnormal blood  
XX pressure, migraine, psychological disorders, respiratory disease, and  
XX coronary heart disease). Histamine may also be used to regulate cellular  
XX growth and tissue repair. The molecules may also be used as components of  
XX vaccines directed against blood-sucking ectoparasites.  
XX  
SQ Sequence 203 AA;  
XX  
Query Match 30.8%; Score 288; DB 20; Length 203;  
Best Local Similarity 35.4%; Pred. No. 8.2e-22;  
Matches 63; Conservative 35; Mismatches 66; Indels 14; Gaps 8;  
OY 1 NOP-----DW-ADEANGAHODAMKSLKADVENYVYKATYKNDPVGNDFTGCVGMAND 55  
DB 27 NEPLKNTWHSKELKN--YDAMWSINOVSTTYFLKSTNNDSWKNNFTCSVTVTS 84  
OY 56 VNEDEKSIQAEFLPMNADTMNOFATEKVTAVKMGYRENAFRETEDEGQVTDVIANSD 115  
DB 85 KHESTFTEYENTYTKKNSQOQVSM-TEYNTAVOEGSVYKKNITQTTNNKKNFNTVFT 143  
OY 116 D-DNCQDVIVPPTDGNBEGYELMT--TDY-DNIPANCLKNFENYAVGRETRDVPFSA 170  
DB 144 DGQTCDLXLYIP-----YKENGVELMWRSDYLQNTPTCCQFIFDLVALGRTTYNISPPDCV 198





**THIS PAGE BLANK (USPTO)**





RL Arch. Pharm. Res. 24:316-322(2001).  
 DR EMBL: AY029222; AAK31641.1; -  
 KM Transferase.  
 SO SEQUENCE 598 AA; 66501 MW; 4CE9BB26545F62A5 CRC64;

Query Match 7.7%; Score 87; DB 2; Length 598;  
 Best Local Similarity 24.0%; Pred. No. 28;  
 Matches 46; Conservative 25; Mismatches 55; Indels 66; Gaps 10;

OY 27 AGENPMAHEE-----LGGKYDAMKSIDOQSVTVYVLAHTYEN-----66  
 DB 377 AGNN--WAHVNSTAYAKDDSIILSSRHGCVN-IGRDQVKNWILAPSGMKNALSKLL 433  
 OY 67 -DTGSGWGFCKQVOEIERKEDYTVT-----SVTFERN-----ASSPIK 106  
 DB 434 KPYDDGKGNALCKDCKNCEKTDFTYQHTAMLSKGLTITFDNGDGLQEPALPTK 493  
 OY 107 Y-----YVWTE--TVKAVFOYG-----YKNIRNAIEYO-----VGGGLNTPTL 143  
 DB 494 YSHFVEYKIDKKGTQVQVWEYGERGYDFYSPITSVIEYQKDRDTMFGGGINLFDVG 553  
 OY 144 IFPDGELDVEY 155  
 DB 554 OPTIGKINETDY 565

## RESULT 9

OY 092N47 PRELIMINARY; PRT: 932 AA.

AC 092N47;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Putative outer membrane receptor protein.  
 GN R02383 OR SMC02721.  
 OS Rhizobium melioli (Sinorhizobium melioli).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1021.  
 RX MEDLINE-21396507; PubMed-11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Mesny D.,  
 RA Pohl T., Portetelle P., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium melioli strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL: AL591790: CAC46962.1; -  
 DR InterPro: IPR000531; TonB\_boxC.  
 DR Pfam: PF00593; TonB\_dep\_Rec; 1.  
 KW complete proteome.  
 SO SEQUENCE 932 AA; 101872 MW; E4F6473BAF5AC484 CRC64;

Query Match 7.7%; Score 86.5; DB 16; Length 932;  
 Best Local Similarity 25.5%; Pred. No. 55;  
 Matches 41; Conservative 24; Mismatches 47; Indels 49; Gaps 11;

OY 44 DANKSIDOQSVTVYVLAHTYENDTG---SMGSO---FKCLQ--VOEIERKEDYTVTSVF 96  
 DB 221 DGRKIDRG-----GDTGFDLIDMGKLGFRADSTQNFREFYQVYDETSNE 267  
 OY 97 TERNAS-----SPIRYN-----VETKAVFOYGYKKNIRNAIEYVGGGLNTPTLIF 145  
 DB 268 TYGLTLIDFKDPPFRYRNASQKVIDAAMNIYOFSH-----STEIPDINLTTTVYR 320  
 OY 146 TDG-----ELCDVEYVPMADGCELMVKKSHYKHPD-YCT 180  
 DB 321 TDTTRAMVKNLNDV--RNDSDTG--WVLSNIIADPOTYST 356

## RESULT 10

OY 08XMD0 PRELIMINARY; PRT: 399 AA.

AC 08XMD0;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein CPE0759.  
 GN CPE0759.  
 OS Clostridium perfringens.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-13 / Type A;  
 RX PubMed-11792842;  
 RA Shimizu T., Ohtani K., Hiraoka H., Ohshima K., Yamashita A.,  
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;  
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
 RT flesh-eater".  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
 DR EMBL: AP003188; BAB80465.1; -  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 399 AA; 46802 MW; 61019F0075828815 CRC64;

Query Match 7.6%; Score 86; DB 16; Length 399;  
 Best Local Similarity 23.8%; Pred. No. 20;  
 Matches 30; Conservative 21; Mismatches 37; Indels 38; Gaps 4;

OY 42 YODAMS---IDCG-----VSVTYVLAHTYENDTGSGWGFCKQVOEIER 85  
 DB 224 YEEAKSLSEYISGFLMCKRSYSENTEMLSEFRISTNSIKASKYKFAKMSLMD-- 281  
 OY 86 KEDDYTVSTFFERNASPEIKYVNV-----TEVKAVFOYGYKKNIRNAIEYVGGGL 137  
 DB 282 -----YDNSKIYKINELGIIIRLFELAEIKVHOYCEYENIGPILEYDKSHGM 329  
 OY 138 NITDTL 143  
 DB 330 NLELTL 335

## RESULT 11

OY 051717 PRELIMINARY; PRT: 584 AA.

AC 051717;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical 63.2 kDa protein precursor.  
 OS Pseudomonas fluorescens.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 7139;  
 RX MEDLINE-93050280; PubMed-7961472;  
 RA Schlatter A., Jendrossek D.;  
 RT "Molecular characterization of the extracellular poly(3-  
 RT hydroxyoctanoic acid) [P(3HO)] depolymerase gene of Pseudomonas  
 RT fluorescens GSK3 and of its gene product.";  
 RL J. Bacteriol. 176:7065-7073(1994).  
 DR EMBL: U10470; AA64539.1; -  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 26  
 SO SEQUENCE 584 AA; 63249 MW; 3AA16FAE898D2CBF CRC64;

Query Match 7.6%; Score 86; DB 2; Length 584;  
 Best Local Similarity 27.1%; Pred. No. 33;  
 Matches 29; Conservative 15; Mismatches 51; Indels 12; Gaps 2;

RP SEQUENCE FROM N.A.  
 RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
 RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
 RA Mather T.N., Ribeiro J.M.C.;  
 RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
 RT scapularis.";  
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF483718; AAM3640.1;  
 SQ SEQUENCE 311 AA; 36581 MW; D7E529FEFE4CFA9A CRC64;

Query Match 9.2%; Score 103; DB 5; Length 311;  
 Best Local Similarity 19.6%; Pred. No. 0.45;  
 Matches 59; Conservative 35; Mismatches 79; Indels 128; Gaps 12;

OY 12 VSAALATQAEFTTSAKAGENPLMAHEELGKYODAKSIDGVSVTYLAKTYENDPTGSW 71  
 DB 13 VSAALAEVEFOSWDAPNPNPDLNKRDJGAMQDARTTKFTANHSYLYLSS-----GW 66  
 OY 72 GSQ-----FKCLOVQEIREEEDYVTSVTFERNASSPIKYYN-----VTEYKAVF 118  
 DB 67 GTRHEIVEDVRCLOVHSSDLN-----YTLKSNYTSKMYNRTSRKMSNSTOYVQAK 117  
 OY 119 Q-YGKNIK-----NAIEQV-----GGGLNI-----TDFLI 144  
 DB 118 QKYSIENIMHIGOPQREVTSPNGTCYMLNFILCESGCRHHEQECWKRTKSEKYV 177  
 OY 145 FTDGELGVFYVPNADG---CELM-----YKSH----- 171  
 DB 178 LFTSPLCYVNSLDDDEYSECEFLSDWIKNTTIPQVVTLEKEDSEDESEEREE 237  
 OY 172 -----YKHVPDYCTEYFNVECA-----KDRKTYDFN 198  
 DB 238 SYERKESVQCEEPRTANKTFTLYDLLEKELPSSCRVAFNLNCGPKRYRDKDCDDIN 297  
 OY 199 E 199  
 DB 298 E 298

## RESULT 6

O95WZ5 PRELIMINARY; PRT; 221 AA.  
 AC O95WZ5;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 OX NCBI\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.;  
 RT "Salp25D, an Ixodes scapularis antioxidant, is one of 14  
 RT immunodominant antigens in engorged tick salivary glands.";  
 RL J. Infect. Dis. 184:0-0(2001).  
 DR EMBL: AF209913; AAK97816.1;  
 SQ SEQUENCE 221 AA; 25359 MW; 7B995AC9C0AFA252 CRC64;

Query Match 8.7%; Score 97.5; DB 5; Length 221;  
 Best Local Similarity 23.4%; Pred. No. 0.89;  
 Matches 49; Conservative 43; Mismatches 80; Indels 37; Gaps 14;

OY 10 TFVSAALATQAEFTTSAKAGENPLMAHEELGKY--ODAKSIDGVSVTYVLA 60  
 DB 31 TYGSTGTTTTCGTGAGATVTTAPEDP-----SKYEQNATRYVE--MNAIOWK 80  
 OY 61 KTVYENDTGSWGSQFKCLOVQEIREEEDYVTSV-FTERNASSPIKYYNTET--YKAV 117  
 DB 81 WRYYD-VTDFSGNVPQCEFNFRVMEKR-----TPTNVSFGYRYRKSN-SMETIDETLLIKDI 134

OY 118 FOYGNIRNAIEYVGGLNITDPLI-FTDGELOVFPVNADG---CELMYK--SHY 172  
 DB 135 GEHPNPVNMFORPPIG---TATDNLVYSNVCTVLRIPFTNGERHCDLMANLPLS 191  
 OY 173 KHVDPCTEYFNVEFCADKRTYDIFNEEC 201  
 DB 192 QETPDDCLNKFEEVC-NTTQIYRYVYVPC 219

## RESULT 7

O8IDR7 PRELIMINARY; PRT; 972 AA.  
 AC O8IDR7;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)  
 DE Hypothetical protein.  
 GN PF13\_0230.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,  
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,  
 RA Submitted (Sep-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL844509; CAD52552.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 972 AA; 119242 MW; 70471A6C0327C32 CRC64;

Query Match 7.9%; Score 89; DB 5; Length 972;  
 Best Local Similarity 19.8%; Pred. No. 35;  
 Matches 43; Conservative 33; Mismatches 81; Indels 60; Gaps 9;

OY 4 QVLLTFVSAALATQAEFTTSAKAGENPLMAHEELGKYODAKSIDGV-----SVYV 58  
 DB 289 KICLFIYISCSYISQYIKHMKENN-----VEKKRKMKSIDNFIKNIINECKHI 340  
 OY 59 LAKTYENDTGSWGSQFKCLOVQEI---RKEEDYVTSVTFERNASSPIKYYNTETV 114  
 DB 341 LMDNLIN-----KCINDQYRITILARKKCSYMLNIVLAKYLSIKYILNINIK 390  
 OY 115 KAVQ--YGKN---IRAIYQVGGCLNITDPLIFDGLCD-VFVVPNADGCELMVK 168  
 DB 391 KSFYICLYLYNSYIFNKSIY-----CGLIFL-----LQKY 423  
 OY 169 KSHYKHVPDYCTEYFNVEFCADKRTYDIFNECYVNG 205  
 DB 424 NLHTYLYQILLIFNPLKDKRKYIDYKHHLLYG 460

## RESULT 8

O939L2 PRELIMINARY; PRT; 598 AA.  
 AC O939L2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Arylsulfate sulfotransferase.  
 GN ASTA.  
 OS Citrobacter freundlii.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Citrobacter.  
 OX NCBI\_TaxID=546;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-M84-8242;  
 RC MEDLINE=21425731; PubMed=11534764;  
 RX Kang J.W., Jeong Y.J., Kwon A.R., Yun H.J., Kim D.H., Choi E.C.;  
 RA "Cloning, sequence analysis, and characterization of the asta gene  
 RT encoding an arylsulfate sulfotransferase from Citrobacter freundlii.";



OY 181 FVFNVCADKRTYDIFNEECVYNGEPM 209  
 DB 181 FVFNVCADKRTYDIFNEECVYNGEPM 209

## RESULT 2

OY 08MW98 PRELIMINARY; PRT; 210 AA.

AC 08MW98; 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DE Putative 22.5 kDa secreted protein.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 OX NCBI\_Taxid=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
 RA Valenzuela J.G., Francischelli I.M., Pham V.M., Garfield M.,  
 Mather T.N., Ribeiro J.M.C.;  
 RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
 RT scapularis."  
 RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL:AF483742; AM93664.1; -  
 SQ SEQUENCE 210 AA; 24689 MW; C5B75614649A9BED CRC64;

Query Match 11.2%; Score 126; DB 5; Length 210;  
 Best Local Similarity 24.3%; Pred. No. 0.0025;  
 Matches 53; Conservative 41; Mismatches 90; Indels 34; Gaps 12;

OY 6 VLLTFVSAALATQAEFTSAKAGENPLMAHELLGKYODAMKSIDGVSVTYVLAITYE 65  
 DB 3 VMHFSFLCLLAILVD---AKGEIIRI-DEDEKYMOYODIORALNPDRESMLYR-TYR 57  
 OY 66 NIDGSGSOFKCL-OVOEIERKEEDYTVTSVTFPNMNASPIKYYVTEYKAVFQYGR 123  
 DB 58 REDD-GSEHICSAVASENPNQSDYEVQERLCT-----KEQTRKTV-TLVATPYK 109  
 OY 124 NIRMALRYGVGGGLNTD-----TLIFTDELCDVFYV--PNADOGCELAV-KKS 170  
 DB 110 TENHATQORNNMARMRSQKKADQDGKGLIYSDVAKCILRLRNSGHDCGLYLSKA 169  
 OY 171 HKHNPDYCTFYFNVCADKRTYDIFNEECVYNGEPM 208  
 DB 170 LDDGVRECEVYGIACGKDEPSY---KORVY-YPM 201

## RESULT 3

OY 08MW98 PRELIMINARY; PRT; 306 AA.

AC 08MW98; 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DE Putative secreted histamine binding protein.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 OX NCBI\_Taxid=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
 RA Valenzuela J.G., Francischelli I.M., Pham V.M., Garfield M.,  
 Mather T.N., Ribeiro J.M.C.;  
 RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
 RT scapularis."  
 RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL:AF483742; AM93664.1; -  
 SQ SEQUENCE 306 AA; 34232 MW; B2DB09181CC4395D CRC64;

Query Match 10.4%; Score 117.5; DB 5; Length 306;

Best Local Similarity 20.0%; Pred. No. 0.023;  
 Matches 46; Conservative 34; Mismatches 73; Indels 77; Gaps 7;

OY 1 MKQVVLTFVSAALATQAEFTSAKAGENPLMAHELLGKYODAMKSIDGVSVTYVLA 60  
 DB 1 MGLMYAALFACIASASAMQAMTIRPENRPL-LNPSIGPLQSAMKAIKSKKDPFLVM 59  
 OY 61 -----KTYENDGSGSOFKCLQVOEIERKEEDYT 91  
 DB 60 FRSRHNHPNITCVVYVYASLHNETLKIYNFTFTYNET--DCKNDPLEYQVALNQTQYK 116  
 OY 92 VNSV-----FFRNASSPIKYYNTEYKAVFOYGYKNI 125  
 DB 117 LENVIRAGIKTPSDKPPPLGSMNYEGDYSCNTSSKPL-----TDMKLA----- 163  
 OY 126 RNAIEYOVG-----GLNITDPLIFTDELCDVFYVFNADOGCELAVKKS 170  
 DB 164 KDAVSEAGSAPFVEGVNLYDFYVHNOHCHILRSLPLKGGCDFLWRS 213

## RESULT 4

OY 0819T9 PRELIMINARY; PRT; 176 AA.

AC 0819T9; 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DE Putative 23 kDa secreted protein.  
 OS Ornithodoros savignyi.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Argasidae; Ornithodoros.  
 OX NCBI\_Taxid=69826;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Salivary gland;  
 RA MEDLINE-21317941; PubMed-11425229;  
 RX Mans B.J., Venter J.D., Vrey P.J., Louw A.I., Nelitz A.W.;  
 RT "Identification of putative proteins involved in granule biogenesis of  
 RT tick salivary glands."  
 RL Electrophoresis 22:1739-1746(2001).  
 DR EMBL:AF452891; AAN76831.1; -  
 SQ SEQUENCE 176 AA; 19375 MW; 44B7595F3BCB6A25 CRC64;

Query Match 9.9%; Score 111.5; DB 5; Length 176;  
 Best Local Similarity 19.2%; Pred. No. 0.038;  
 Matches 30; Conservative 33; Mismatches 64; Indels 29; Gaps 6;

OY 44 DAMKSIDGVSVTYVLAITYENDGSGSOFKCLQVOEIERKEEDYTVTSVTFPNMNAS 103  
 DB 23 DVM-NVLKGSDSKFLMVKRTTERGAN-----KCYMKRTSMDESSHTLEVLGYSKAGT 75  
 OY 104 PIKY-----YNYTEYKAVFOYGYKNIR-----NAIEYOVGGGLNITDPLIFTDELCDV 153  
 DB 76 TDFVPSKSYTYVYATSEGAASYNNMVTYRGASHGKFE-----LVYSDDGCGNI 125  
 OY 154 FYVPNA--DQGEIYKSKHYKHPDYCTCFVNVFC 187  
 DB 126 LQMKTSFPFGKCELMAPBEKARNVSESSCGKFEKLC 161

## RESULT 5

OY 08MW98 PRELIMINARY; PRT; 311 AA.

AC 08MW98; 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DE Putative secreted histamine binding protein.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 OX NCBI\_Taxid=6945;  
 RN [1]



Mon Aug 4 17:50:34 2003

dret6.rsp

Page 10

Search completed: August 4, 2003, 17:15:10  
Job time : 25 secs

---

```

OY 81 OEIERKEEDYVTSVTEFRNASSPIKYYNVTETVKAVFOYGYKNINNAIEYOYGGGLNT 140
DB 210 -----KGEUMALTAGHROKHAKKVKYREFPASTSGVCWNPIDELRLGTEYEGDVONIA 264
OY 141 DFLIFDGLCDVDFYV-----PNADGCE 164
DB 265 TLIVDPDGKGLDSHMQKTAFFALLVGYLHLYAKADGGTATLPSVDANLADPNRIG-E 323
OY 165 LMKVKSRYKHV 175
DB 324 LWMEMATYGHV 334

RESULT 14
DNK_RICPR STANDARD; PRT; 627 AA.
AC Q9ZDX9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein dnk (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70).
GN DNK OR Rp185.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_Taxid=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Neeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of mitochondria."
RL Nature 396:133-140(1998).
CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -1- INDUCTION: By stress conditions e.g. heat shock (by similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ235270; CAA14651.1;
CC PIR: D71729; D71729.
CC HSP: P04475; 1D64.
DR HAMAP; ME_00332; 1.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PRO0301; HEATSHOCK70.
DR PRODOM; PD000089; HSP70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation;
KW Complete proteome.
FT MOD_RES 197 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 627 AA; 68383 MW; A638A0A5A3C8AC7 CRC64;

Query Match 7.2%; Score 81.5; DB 1; Length 627;
Best Local Similarity 28.2%; Pred. No. 12;
Matches 33; Conservative 15; Mismatches 42; Indels 27; Gaps 7;

OY 72 GSOEKLQVOE-----IERKEEDYVTSVTEFRNA-----SSPIRYNVTETVKAVFOYG 121
DB 10 GTTNCVAVMEGKEPKVIDNAEGERTTSIIAFANSERLVGPAPKRAVTV----- 59

```

```

OY 122 YKNINNAIEYOYGG--GLNITDPLIFDGLCDVDFYVPNADGCELMVKSHYKHP 176
DB 60 --NPRNTI-YAVKRLIGRNFETDPMVKRQGLVP-INIYKADNG-DAWVEADNHYKSP 111

RESULT 15
HS9C_DICDI STANDARD; PRT; 700 AA.
ID HS9C_DICDI
AC P54651;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Heat shock cognate 90 kDa protein.
GN HSPD OR HSC90.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostellium.
OX NCBI_Taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3.
RA Boves J., Dittlich W., Minter U., Lottspeich F., Gerisch G.,
RA Falx J.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L43591; AAA69917.1;
CC HSP: P07900; 1YER.
DR Dictydb; DD00051; hspd.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR001404; HSP90.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00183; HSP90; 1.
DR PRINTS; PRO0775; HSPATPase_C; 1.
DR SMART; SM00387; HATPase_C; 1.
DR PROSITE; PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock.
SQ SEQUENCE 700 AA; 79881 MW; BFB5ECF238089CD3 CRC64;

Query Match 7.2%; Score 81; DB 1; Length 700;
Best Local Similarity 19.5%; Pred. No. 15;
Matches 52; Conservative 32; Mismatches 105; Indels 78; Gaps 9;

OY 11 FVSALATQAEITTSKAG--ENPLWA-----HEELGK-----Y 42
DB 128 FYSAYLVADYIVHKNDDDEQYVWESSAGGEFTALDHTPELGRTIVLHMKEDQDLY 187
OY 43 ODAMK-----SIQGVSVTVYLAHTTYENDTSGWSGQFC-LOVOIERKEEDYT 91
DB 188 LDETQIKNLVKKHSEFIQYPISLTIKKEKEDDEETAKEGEESTDAKIEIEEKEKK 247
OY 92 V---TSVTFRNASSPIKYYNVTETVKAVFOYGYKNINNAIE-----YOYGGGLNTD 141
DB 248 VKVOEKEDVNLKTKPLTRNPSDVTKEEYNSFYKSIISNDWEPLAVKHSVEGOLERKA 307
OY 142 TLIFDGLCDVDFYVPNADGCELMVKSHYKH-----VPDYCTFEVNFCAKD----- 190
DB 308 ILFVPKKAPFLDFESKKANNIKLVKRVFTLMDNCADLIIPYLNFRGIVDSEDLPNIS 367
OY 191 -----KTYDIFNE 199
DB 368 RETLQKNILTVIRKNLVKKCIELFNE 394

```

```

Oy      124 NRNAIEYGVGGILNTDTLFTDGLCDVYVVPNADGCELW 166
      113 NLRNIT-----GDLELDQTL--TSREMINIQLREVDEARDKN 148

RESULT 12
RIR2_CABEL
ID RIR2_CABEL STANDARD; PRT; 381 AA.
AC PA2170;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
GN (Ribonucleoside reductase).
OS RNR-2 OR C03C10.3.
OC Caenothabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderridae; Caenothabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Becky M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioresoxin + H(2)O -> ribonucleoside diphosphate + reduced
CC thioresoxin.
CC -1- COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z35637; CA84688.1; -.
CC PIR; T18876; T18876.
CC HSSP; P1157; IXSM.
CC Wormpep; C03C10.3; CE00874.
CC InterPro; IPR000358; RibonucL_reductase.
CC Pfam; PF00268; Ribonuc_red_sm; 1.
CC PROSITE; PS00368; RIBORED_SMALL; 1.
CC Oxidoreductase; DNA replication; Iron.
FT METAL 130 130 IRON 1 (BY SIMILARITY).
FT METAL 161 161 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 164 164 IRON 1 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT METAL 258 258 IRON 2 (BY SIMILARITY).
FT METAL 261 261 IRON 2 (BY SIMILARITY).
FT ACT_SITE 168 168 BY SIMILARITY.
SQ SEQUENCE 381 AA; 44289 MW; 75497147ABF36C9 CRC64;

Query Match 7.3%; Score 82; DB 1; Length 381;
Best Local Similarity 21.8%; Pred. No. 5.9;
Matches 29; Conservative 20; Mismatches 56; Indels 28; Gaps 3;
Oy      9 LTFVSALATQAETTSKAGENPLMAHEELG-----KYDAWKSSIOGVSVTY 57
      35 LEIVDOTKAASAEETNSEVENEDADEPMLQDLNRFVYPLKHHIDIMFYKKAVASFW 94
Oy      58 VLAKTTYENDTSGSQFKLOVLEIKKEDYTVTSVFTFRNASSPI-----KYYNV 110
      95 TVEEVDGKDMNDW-----EKNNGDEQYFISILAFPAASDSIVENICERSNE 144
Oy      111 TETVKAIFYGYGK 123

```

```

Db      145 VQVSEARFFYGFQ 157
      145 VQVSEARFFYGFQ 157

RESULT 13
TRG5_ECOLI
ID TRG5_ECOLI STANDARD; PRT; 637 AA.
AC O00184;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Conjugal transfer protein trag.
GN TRAG.
OS Escherichia coli.
OC Plasmid Incp-beta R751.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HB101;
RX MEDLINE-92190548; PubMed-1665997;
RA Ziegelin G., Pansegrau W., Streck B., Balzer D., Kroeger M.,
RA Kruff V., Lanka E.;
RT "Nucleotide sequence and organization of genes flanking the transfer
RT origin of promiscuous plasmid RP4."
RL DNA Seq. 1:303-327(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Thomas C.M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Required for conjugal transfer of plasmid R751.
CC Binds tightly and specifically to the relaxase tral. Can also bind
CC to DNA without sequence specificity. May form a pore-like
CC structure that could serve as a channel for DNA transfer (By
CC similarity).
CC -1- SUBUNIT: May form multimers of at least 18 subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE VIRD4/TRAG FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X54458; CA38327.1; -.
CC EMBL; U67194; AAC64474.1; -.
CC InterPro; IPR003688; TRAG.
CC Pfam; PF02534; TRAG; 1.
CC Plasmid; Conjugation; DNA-binding; Transmembrane; Inner membrane.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 43 POTENTIAL.
FT DOMAIN 44 84 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 85 105 POTENTIAL.
FT DOMAIN 106 637 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 637 AA; 69883 MW; 7B45A9A9020902FB CRC64;

Query Match 7.3%; Score 82; DB 1; Length 637;
Best Local Similarity 18.3%; Pred. No. 11;
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;
Oy      5 VLLTFVSALATQAETTSKAGF---NPLMAHEE----- 37
      90 LVSTVGILGVAAVAKVYTSNSSKANLEYLHGSARWAEKKDQAGLPRERNVLEIVYKAA 149
Oy      38 --LLGKYDAWKSSID-----QGVSVTVYLAKTYYENDTG-----SMGQFKCLOY 80
      150 PRTAGVYVGGMQDKGNFTYLRHSGPEHVLTATPRTSGKGVGLVPTLLSGASSVITDL 209

```

```

OY 101 ASPTKRYNVEETKAVQYCKNLRN-----IEYGGGINTDITLF 145
DB 252 SMKALMTKSKREVSDEYKEEYKHLAHAWDDPLEYIANKACSTFYQ-----ALLF 302
OY 146 TDGELCDVEYVPNADQCELMVKKSHY-----KHVPDCTEFVNFVCARD 190
DB 303 IPSHAFDLEFNSDAKIGMQLYVKRFVIMSDCDQLMPYLRFRKGVVDAED 352

RESULT 10
ID HTPG_RICPR STANDARD; PRG; 621 AA.
AC Q9ZCB9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature protein G).
GN HTPG OR RP840.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RA MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Slicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RL mitochondria."
RL Nature 396:133-140(1998).
CC -1 FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -1 SUBUNIT: Homodimer (By similarity).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1 SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ235273; CAA15264.1; -.
DR PIR: H71645; H71645.
DR HSSP: P02829; 1A4H.
DR HAMAP: MF_00505; 1.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR001404; Hsp90.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00183; HSP90; 2.
DR PRINTS: PRO0775; HEATSHOCK90.
DR SMART: SM00387; HATPase_C; 1.
DR PROSITE: PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 328 A; SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 329 544 B (BY SIMILARITY).
FT DOMAIN 545 621 C.
SQ SEQUENCE 621 AA; 70713 MW; 9F97B97801524007 CRC64;

Query Match 7.5%; Score 84.5; DB 1; Length 621;
Best Local Similarity 18.9%; Pred. No. 6.4;
Matches 39; Conservative 42; Mismatches 88; Indels 37; Gaps 8;

11 FVSAALTOAEETTSKAKGNIAMHE-ELGKY--DDAKMSIDGVSITYVLAK--ITYE 65
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 YSSFMADKVTYSRRKGESVHTWESDGLSEYIVAAOSEGFTGTELVLYIKKSETTEL 186

```

```

OY 66 NDTGSMGSGFCLOAQOETREKEDYTVTSVTFERNASSPIKYYN-----TETKAVFOY 120
Db 187 DH-----FRKATKIVKSYSDHIANVPIFCDEAGNNEIQANSASALMTRPKSEIED 236
OY 121 GYKNIRNAIEYQVG-----GGINTDTLTFTDGLDVEFYVBNADQCELMVWK 169
Db 237 QYKEFYKSLSYAVDDPPWTLTNNKNGEALIEFTNLTLPSSKTFDLFH--DDRKRKYVLTIKR 295
OY 170 -----SHKHVPDYCTFVFNFCARD 190
Db 296 VFISDENIDLPSYLRFLRGVDSED 321

RESULT 11
ID Y692_METTH STANDARD; PRT; 318 AA.
AC 026788;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MTH692.
GN MTH692.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=Delta H;
RC MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubots J.,
RA Alredge T., Bashlitzaden R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumu M., Potlier B., Qiu D.,
RA Spadafora R., Vlcare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Seifer H., Patwell D., Prabhakar S.,
RA McDonnell S., Shlmer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delta: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AE000848; AAB85197.1; "-
CC DR PIR; C69192; C69192.
CC DR InterPro; IPR001107; Band_7.
CC DR InterPro; IPR001972; Stomatrin.
CC DR Pfam; PF01145; Band_7; 1.
CC DR PRINTS; PR00721; STOMATIN.
CC DR SMART; SM00244; PHB; 1.
CC DR PROSITE; PS01270; BAND_7; 1.
CC KM Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSEM 22 POTENTIAL.
CC SQ SEQUENCE 318 AA; 35425 MW; 4472C0AC9E9927BD CRC64;

Query Match 7.3%; Score 82.5; DB 1; Length 318;
Best Local Similarity 23.3%; Pred. No. 4.3;
Matches 38; Conservative 23; Mismatches 77; Indels 25; Gaps 6;

OY 5 VLLITFVSALATQAEFTSAKAGENPLMAHELLGKYQDAWKSIIDGVSVTVYLAQTT 64
Db 10 VLLVIAKSKILIRPEKGVV-----ERLGRKYQ--RTVESGLVLIIFIEIAIK 55
OY 65 ENDTGSMGSGFCLOAQOETREKEDYTVTSVTFERNASSPIK--YYNTETKAVFOYGYK 123
Db 56 KVD--NRQYVDVDPPEQVITKNTVVVVDVCIIEYVVDPPNAVNAVVDYQATTKLAQT 112

```

```

CC Name=2;
CC IsoId=P55824-2; Sequence=VSP_005270;
CC Name=3;
CC IsoId=P55824-3; Sequence=VSP_005269;
CC -1- TISSUE SPECIFICITY: EYE DISKS AND OVARIES.
CC -1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
CC -1- SIMILARITY: Belongs to peptidase family C19.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04959; AAF01345.1; -
DR EMBL; L04958; AAF01346.1; -
DR EMBL; L04960; AAF01347.1; -
DR EMBL; L04960; AAF01348.1; -
DR EMBL; AE003779; AAF57198.1; -
DR EMBL; AE003779; AAN14291.1; -
DR EMBL; AF145677; AAD38652.1; -
DR MEROPS; C19.007; -
DR FLYBase; FBgn0005632; faf.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0007349; P:cellularization; IMP.
DR GO; GO:0009795; P:embryonic morphogenesis; IMP.
DR GO; GO:0008683; P:mystery cell fate differentiation (sensu Dr. .; IMP.
DR GO; GO:0007097; P:nuclear migration; IMP.
DR GO; GO:0006512; P:ubiquitin cycle; IGI.
DR InterPro; IPR001394; UCH-2.
DR Pfam; PF00443; UCH; 1.
DR PROSITE; PS00972; UCH_2.1; 1.
DR PROSITE; PS00973; UCH_2.2; 1.
DR PROSITE; PS0235; UCH_2.3; 1.
DR UDI conjugation pathway; Hydrolase; Thiol protease;
DR Developmental protein; Vision; Alternative splicing.
DR ACT_SITE 1677 1677 BY SIMILARITY.
DR ACT_SITE 1978 1978 BY SIMILARITY.
DR ACT_SITE 1986 1986 BY SIMILARITY.
DR VARSPLIC 2705 2778 KCRVITIKLVESKDEEDATTAATTAETVTSPPAATA
ATLEPAGMSELTMTWKNLIISQENPQAKSLQ -> VYRA
NNV (in isoform 3).
/FTId-VSP_005269.
IATAATLEPAGMSELTMTWKNLIISQENPQAKSLQ ->
VARSPLIC 2742 2778 SORROL (in isoform 2).
/FTId-VSP_005270.
CONFLICT 234 234 E -> D (in REF. 1).
FT CONFLICT 2725 2725 T -> S (in REF. 1; AAF01345).
SQ SEQUENCE 2778 AA; 31139 MW; FFB90438BA53A02B CRC64;
Query Match 7.6%; Score 86; DB 1; Length 2778;
Best Local Similarity 22.9%; Pred. No. 29;
Matches 54; Conservative 28; Mismatches 72; Indels 82; Gaps 12;

```

```

RESULT 9
HTPG MYCLE
ID HTPG MYCLE STANDARD: PRT; 656 AA.
AC 033012;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature
DE protein G)
GN HTPG OR ML1623 OR MLCB250.19C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Elgimeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Mole S.,
RA Murphy L., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrall B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z97369; CAB10613.1; -
DR EMBL; AL583922; CAC30574.1; -
DR PIR; A87112; A87112.
DR HSP; P07900; IBYQ.
DR Lepioma; ML1623; -
DR HAMAP; MF_00505; -; 1.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR001404; Hsp90.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00183; HSP90_2.
DR PRINTS; PR00775; HEATSHOCK90.
DR SMART; SM00387; HATPase_c; 1.
DR PROSITE; PS00298; HSP90; 1.
DR Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 359 A: SUBSTATE-BINDING (BY SIMILARITY).
FT DOMAIN 360 575 B (BY SIMILARITY).
FT DOMAIN 576 656 C.
SQ SEQUENCE 656 AA; 73866 MW; A526690CA66E03FF CRC64;
Query Match 7.6%; Score 85; DB 1; Length 656;
Best Local Similarity 21.3%; Pred. No. 6.2;
Matches 49; Conservative 38; Mismatches 79; Indels 64; Gaps 9;

```

entites requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE001265; AAC65938.1; -  
 CC PIR; F71258; F71258.  
 DR HSP; P02839; 1AH8.  
 DR TIGR; TP0964; -  
 DR HAMAP; MF\_00505; -; 1.  
 DR InterPro; IPR003594; ATPbind\_Atpase.  
 DR InterPro; IPR001404; HSP90.  
 DR Pfam; PF02518; HATPase\_c.1.  
 DR Pfam; PF00183; HSP90\_2.  
 DR PRINTS; PR00775; HEATSHOCK90.  
 DR SMART; SM00387; HATPase\_c.1.  
 DR PROSITE; PS00298; HSP90; 1.  
 KM Chaperone; ATP-binding; Heat shock; Complete proteome.  
 FT DOMAIN 1 348 A: SUBSTRATE-BINDING (BY SIMILARITY).  
 FT DOMAIN 349 565 B (BY SIMILARITY).  
 FT DOMAIN 566 639 C.  
 SQ SEQUENCE 639 AA; 72937 MW; 3E8FDBAC2282C31D CRC64;

Query Match 7.7%; Score 87; DB 1; Length 639;  
 Best local similarity 22.9%; Pred. No. 4;  
 Matches 50; Conservative 26; Mismatches 52; Indels 90; Gaps 10;

QY 5 VVLLTFVSALATQAEFTTSARAGENPLMAHEELLGKYQD-----AWKSIDQGV 54  
 DB 182 VVLLHLSQENSEFATR-----WRLEEVIAKKSVDHIAFPVLYHLYQKEDKQGA 228  
 QY 55 VVYVIAKTTYENDTSMGQSFQCLQVQELERKEEDY-----TVTSVFPRNASSPIKYYN 109  
 DB 229 VYDTQKTKDQVNDADALMKRPS-----ELKEEDYHREYQTLT-----RSTPPLLYVH 277  
 QY 110 VTEFKAAYFYQYKRNIRNAIEYQVGGGLNTFTLTFTDQ--ELCDVFYVP-----NA 159  
 DB 278 TK-----ASTGTQYVTLFTVPAKAPDPLDFA 303  
 QY 160 D--QGCELWVK-----SHYKHVPDYCFVFNVFCAD 190  
 DB 304 DYKPGVKLFVKRRVFTDDEKELLPYLRFRGVIDSED 341

RESULT 8  
 FAF\_DROME STANDARD; PRT; 2778 AA.  
 AC P55824; Q9Y9T6; Q9Y027;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)  
 DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease FAF) (Deubiquitinating enzyme FAF) (Pat facets protein).  
 GN FAF OR BCDA; LD2582 OR CG1945.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriidae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND TISSUE SPECIFICITY.  
 RC TISSUE=Eye, Imaginal disk;  
 RX MEDLINE=93202020; PubMed=1295747;  
 RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;  
 RT "The fat facets gene is required for Drosophila eye and embryo development.";  
 RL Development 116:985-1000(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Abmayant A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
 RA Ballew R.M., Basu P.V., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brötter P.,  
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hardek N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibbegan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paclet J.M.,  
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP REVISIONS, AND ALTERNATIVE SPLICING.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Dysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [4]  
 RP SEQUENCE OF 1089-2778 FROM N.A. (ISOFORM 1).  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196012; PubMed=10731138;  
 RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Erise E.,  
 RA Stapleton M., Harvey D.A.;  
 RT "A Drosophila complementary DNA resource.";  
 RL Science 287:2222-2224(2000).  
 RN [5]  
 RP FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A  
 RP ROLE IN COMPOUND EYE ASSEMBLY AND OOGENESIS RESPECTIVELY. IN THE  
 RP LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS  
 RP PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY  
 RP CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR  
 RP NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND  
 RP COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR  
 RP FUNCTION.  
 CC -! CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =  
 CC ubiquitin + a thiol.  
 CC -! ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing: Named isoforms=3;  
 CC Comment-Experimental confirmation may be lacking for some  
 CC isoforms;  
 CC Name=1;  
 CC IsoId=P55824-1; Sequence=Displayed;



```

Db      183 DT--FLDHRP---LKHIVKSYSDHIAVPIFFEDAGNNEIQLNSASALWTRPKSEITEEQ 237
OY      118 FOVCYNININATE-----YQVGGGLNTDITLFIIDGELCDVFPVFNADGCELVK- 169
Db      238 YKEEYFSLSYAIDDPDPIITMHNKKEGAIEETNLNLFIPSSKTFDLEH-PDRKRRVLYIKRV 296
OY      170 ----SHYKHPDYCTFEVNFVFCARD 190
Db      297 FIDENIDILPISRLRLRCVDSED 321

RESULT 6
HTPG_BORBU STANDARD; PRT; 616 AA.
ID HTPG_BORBU
AC P42555;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chapterone protein htpg (Heat shock protein htpg) (High temperature
DE protein G).
GN HTPG OR BB0560
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxId=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-297;
RA Porcella S.F., Radolf J.D., Norgard M.V.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE-98065943; PubMed-9403685;
RA Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwin M.,
RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervilavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
RA Uterback T., Matthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.
RL Nature 390:580-586(1997).
RN [3]
RP SEQUENCE OF 497-574 FROM N.A.
RC STRAIN-212;
RX MEDLINE-9511614; PubMed-7812434;
RA Ojalil C., Davidson B.E., Saint-Girons I., Old I.G.;
RT Conservation of gene arrangement and an unusual organization of rRNA
RT genes in the linear chromosomes of the Lyme disease spirochaetes
RT Borrelia burgdorferi, B. garinii and B. afzelii.
RL Microbiology 140:2931-2940(1994).
CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U51878; AAA97469.1; -
CC DR EMBL: AE001157; AAC66919.1; ALT_INIT.
CC DR EMBL: L32145; AAC41403.1; -
CC DR HSP: P07900; 1YER.
CC DR TIGR: BB0560; -

```

```

DR HAMAP: ME_00505; -; 1.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR001404; Hsp90.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00183; HSP90; 2.
DR PRINTS: PR00775; HEATSHOCK90.
DR SMART: SM00387; HATPase_C; 1.
DR PROSITE: PS00298; HSP90; 1.
KW Chapterone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 333 A: SUBSTRATE-BINDING (By similarity).
FT DOMAIN 334 542 B (By similarity).
FT DOMAIN 543 616 C.
SQ SEQUENCE 616 AA; 71218 MW; D676552F48DBEE84 CRC64;

Query Match      8.5%; Score 96; DB 1; Length 616;
Best Local Similarity 21.5%; Pred. NO. 0.61;
Matches 46; Conservative 34; Mismatches 92; Indels 42; Gaps 8;

OY      11 FVSALATQ-AETTSKAGENP--LWAHEELGKYODANKSIDGVSVTVYLAQTTEND 67
Db      121 FVSATFVSEKVEYTSKKALESADAYIMSDGKGYELEKAKKESGTEIKLYLNKGELE-- 178
OY      68 TGSWSGQFKCLOV-----QETKRKEEDYTVTSVFTFRRNA 101
Db      179 ---YANKMKIOEIKKYSNHNIPYIKYSEPIMKDKGKEIEKEKINETTLMTKN- 234
OY      102 SSPIKYVNTFETVKAFOYGKKNIRNAIEYQGGGLNTDITLFIIDGELCDVFPVFNADQ 161
Db      235 KSEIKAEENYERKNT-TDYENPLMHITTKAEGLNLEYTNLVVSKAPDYLY-PNTRP 292
OY      162 GCELVKVK-----SHYKHPDYCTFEVNFVFCARD 190
Db      293 GVKLFNIRFIIDSEGLLPVNLRFIKGIIDCD 326

RESULT 7
HTPG_TREPA STANDARD; PRT; 639 AA.
ID HTPG_TREPA
AC O83949;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chapterone protein htpg (Heat shock protein htpg) (High temperature
DE protein G).
GN HTPG OR TP0984.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxId=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nichols;
RX MEDLINE-98332770; PubMed-9665876;
RA Frazer C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.
RL Science 281:375-388(1998).
CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

FT CHAIN 19 190 FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN
FT DISULFID 66 187 1.
FT DISULFID 137 166 BY SIMILARITY.
SQ SEQUENCE 190 AA: 21370 MW: 855BE151A9003B1 CRC64:

Query Match
Best Local Similarity 22.4%; Score 252; DB 1; Length 190;
Matches 68; Conservative 28; Mismatches 89; Indels 18; Gaps 8;

OY 3 MOVVLLTFVSAALTAQAEETSAKAGENPLNAHEELLKYDQAKMSIDQGVSYTVYLAKT 62
DB 1 MKLLSLAFVALLSQVKA-----DKPVMADEANGEHODAMKHLQKLYEENDLTKA 52
OY 63 TYENDTGSMSQFCKLQVQEIREEEDYTVSVTFPRNASSPIKYYNTEYKAVFOYGY 122
DB 53 TYKNRP-VWGNDFCTGTAQAQNLNDEKNVEMFEMNADIV-YQHFERKATPPKMGY 110
OY 123 KNIRNAIEYQVGGGLNTDTLFTDGLCDVFEYVNNAD--QGCELMVKKSHYKHPDYC 179
DB 111 -NKENAIYYQFEDGGVLLIDVLAFSQ-DNCYVYIALGPPGSGAGYELMA--TDYTDVPASC 166
OY 180 TFVENVFCAKDRKRTDIFNECV 202
DB 167 LEKENEYAA-GIIPRVDTYTSOCL 188

RESULT 4
NGCG_MESVI STANDARD: PRT; 174 AA.
AC 09MURI;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain J, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain J) (NADH-plastoquinone oxidoreductase
DE: subunit J).
GN NDHJ.
OS Mesostigma viride.
OC Chloroplast.
CC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
CC Mesostigmatales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Lemieux C., Orlis C., Turnel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
RT branch of green plant evolution.";
RL Nature 403:649-652(2000).
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC plastoquinol.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 30 kDa SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF166114; AAF43839.1; -
DR InterPro: IPR001268; Complex1_30k.
DR Pfam: PF00329; complex_30kd; 1.
DR PROSITE: PS001581; Complex1_30k; 1.
DR PROSITE: PS00542; Complex1_30k; 1.
KW Oxidoreductase; NAD: NADP: Quinone; Plastoquinone; Chloroplast.
SQ SEQUENCE 174 AA: 20329 MW: 122ABFA9188D47B7 CRC64:

Query Match
Best Local Similarity 8.7%; Score 97.5; DB 1; Length 174;
Matches 27.0%; Pred. No. 0.097;

```

```

Matches 31; Conservative 20; Mismatches 37; Indels 27; Gaps 5;

OY 105 IKYVNTFVAVAPQYGNKNRMAIEYVGGGLNTDTLFTDGLCDVFEV---PNAD 160
DB 47 VKAOMLVISVAIALYAGFNRLSQCAYDSFG-----GDLASVHLTVDDNAD 95
OY 161 QGCELMVKKSHYKHPDYCTFEVFNVCA-----KDRKTYDIFNEECVNGEPML 209
DB 96 QPQEVCIK-----VFPRPKRPILPSVFWTKRADQERSTYMPG--IYEGSHPL 144

RESULT 5
HTPG_RICCN STANDARD: PRT; 621 AA.
AC HTPG_RICCN
ID P58478;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature
DE protein G).
GN HTPG OR RCL302.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Remesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE008676; AAL03840.1; -
DR PIR: F97862; F97862.
DR HAMAP: MF_00505; -; 1.
DR InterPro: IPR003594; AtPbind_ATPase.
DR InterPro: IPR001404; Hsp90.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00183; HSP90; 2.
DR PRINTS: PR00775; HEATSHOCK90.
DR SMART: SM00387; HATPase_c; 1.
DR PROSITE: PS00298; HSP90.1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 328 A; SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 329 544 B (BY SIMILARITY).
FT DOMAIN 545 621 C.
SQ SEQUENCE 621 AA: 70804 MW: PF6E5B35DDBF017 CRC64:

Query Match
Best Local Similarity 8.6%; Score 96.5; DB 1; Length 621;
Matches 43; Conservative 46; Mismatches 81; Indels 35; Gaps 10;

OY 11 FVSAALTAQAEETSAKAGENP--LMAHEELLGKY--QDAMKSIDQGVSYTVYLAKTYYEN 66
DB 127 YSSFEVADKVTIVTIRKAGEDVHTM-BSDGETVSDSDKEFTRGTEIVLHKV---EE 182
OY 67 DTGSMGSOFKLQVQEIREEEDYTVSVTF-----RNASSPIKYYNTEYKAV 117

```

FT TUR 77 80  
 FT STRAND 81 89  
 FT TUR 91 92  
 FT STRAND 97 106  
 FT TUR 109 110  
 FT STRAND 116 121  
 FT TUR 122 123  
 FT STRAND 126 135  
 FT TUR 136 137  
 FT STRAND 138 143  
 FT TUR 152 157  
 FT STRAND 160 161  
 FT TUR 165 174  
 FT STRAND 175 177  
 FT TUR 181 182  
 FT STRAND 186 188  
 FT TUR 190 AA; 21464 MW; 6923A3E902552B6F CRC64;

Query Match 25.8%; Score 290.5; DB 1; Length 190;  
 Best Local Similarity 36.0%; Pred. No. 8.5e-19;  
 Matches 72; Conservative 32; Mismatches 79; Indels 17; Gaps 9;

Oy 7 LLTFTVSALATQAETTSAGENPLMAHEBLGKYODAKSIDGVSVTVYLAKTYYEN 66  
 Db 3 LLTFTVSALATQAETTSAGENPLMAHEBLGKYODAKSIDGVSVTVYLAKTYYEN 57  
 Oy 67 DTGSGSOPKLOVQELERKEEDYTVSVTF--RNASSPIKYYVTEYKAVPQYKNI 125  
 Db 58 DP-VWGNDFTCVGMANDVDEKSIQAEFLPMNADTNMOF--ATEKTVAKWYG-NR 113  
 Oy 126 RNAIEYOVGGGLNTDTLFTDGLCDVFPVNPAD--OCCELMVKSKHYKHPDYCTFV 182  
 Db 114 ENAFRYETEDGQVETDVIASD-DNCQVIYVPGTDGNECEYELMT--TDYDNPANLKN 170  
 Oy 183 FNVFCADKORTYDIFNEECV 202  
 Db 171 FNEY-AVGRETRDVFSTACL 189

## RESULT 2

HBPM\_RHIAP STANDARD; PRT; 200 AA.  
 ID HBPM\_RHIAP STANDARD; PRT; 200 AA.  
 AC 077422;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Male-specific histamine-binding salivary protein precursor (MS-HBP).  
 OS Rhipicephalus appendiculatus (Brown ear tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.  
 NCBI\_TaxID=34631;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;  
 RT "Tick histamine-binding proteins: Isolation, cloning, and three-dimensional structure".  
 RT Mol. Cell 3:661-671(1999).  
 CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO  
 CC OUTCOMPLETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO  
 CC SUPPRESS INFLAMMATION DURING BLOOD FEEDING.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U96082; AAC63108.1; -.

DR HSSP; 077421; 10FT.  
 DR InterPro: IPR002970; His\_binding.  
 DR Pfam: PF02098; His\_binding; 1.  
 DR ProDom: PD152455; His\_binding; 1.  
 KW Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 200  
 FT DISULFID 65 193  
 FT DISULFID 137 169  
 FT SEQUENCE 200 AA; 22851 MW; C46A1C8C6BCA008 CRC64;

Query Match 23.6%; Score 265; DB 1; Length 200;  
 Best Local Similarity 34.5%; Pred. No. 1.6e-16;  
 Matches 69; Conservative 34; Mismatches 71; Indels 26; Gaps 10;

Oy 3 MOVVLLTFVSALATQAETTSAGENPLMAHEBLGKYODAKSIDGVSVTVYLAKT 62  
 Db 1 MKVLLTV--LGAALCONADA-----NPTWANEAKLGSYDAMKSIDQDNKRRYLAQA 51  
 Oy 63 TYENDGSGSOPKLOV--QELERKEEDYTVSVTFRRAS--SPIKYYVTEYKAVPQ 119  
 Db 52 TQRTD-GWGEETFCVSVNAEKIGKKLNATF---LYKKHLTDLKESETTFWKAVD 106  
 Oy 120 YGKKNRNAIEYOVGGGLNT--DTLFTDGLCDVFPVNPAD--NADG--CELMVKSKHYK 173  
 Db 107 Y---TTENGIKYETGGRITQTFEDVFPFSYKNCQVIFVPEKSGSDGDELMVSEDK 163  
 Oy 174 HVPDYCTFVNVFCADKRT 193  
 Db 164 KIPDCCKFTMAVFAQOQERT 183

## RESULT 3

HBPM\_RHIAP STANDARD; PRT; 190 AA.  
 ID HBPM\_RHIAP STANDARD; PRT; 190 AA.  
 AC 077420;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Female-specific histamine-binding protein 1 precursor (FS-HBP).  
 OS Rhipicephalus appendiculatus (Brown ear tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.  
 NCBI\_TaxID=34631;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;  
 RT "Tick histamine-binding proteins: Isolation, cloning, and three-dimensional structure".  
 RT Mol. Cell 3:661-671(1999).  
 CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO  
 CC OUTCOMPLETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO  
 CC SUPPRESS INFLAMMATION DURING BLOOD FEEDING.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U96080; AAC63106.1; -.  
 DR HSSP; 077421; 10FT.  
 DR InterPro: IPR002970; His\_binding.  
 DR Pfam: PF02098; His\_binding; 1.  
 DR ProDom: PD152455; His\_binding; 1.  
 KW Signal.  
 FT SIGNAL 1 18  
 FT POTENTIAL.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 17:13:07 ; Search time 24 seconds

(without alignments)  
409.524 Million cell updates/sec

Title: DRET6  
Perfect score: 1125  
Sequence: 1 MKMQVLLTFVSALATQA.....DRKTYDIFNECVNGEPM 209

Scoring table: Biosum62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290.5	25.8	190	1	HBP2_RHIAP
2	265	23.6	200	1	HBPM_RHIAP
3	232	22.4	190	1	HBP1_RHIAP
4	97.5	8.7	174	1	NUGC_MESVI
5	96.5	8.6	621	1	HTPG_RICCN
6	96	8.5	616	1	HTPG_BORBU
7	87	7.7	639	1	HTPG_TREPA
8	86	7.6	2778	1	FAR_DROME
9	84.5	7.5	656	1	HTPG_MYCLE
10	82.5	7.3	318	1	Y692_METTH
11	82.5	7.3	381	1	RIR5_CAEEL
12	82	7.3	637	1	TRG5_ECOLI
13	81.5	7.2	627	1	DNAX_RICPR
14	81	7.2	700	1	HS9C_DICDI
15	80.5	7.2	621	1	HTPG_HELPJ
16	79.5	7.1	621	1	HTPG_HELPY
17	79.5	7.1	629	1	HTPG_CHLFE
18	78.5	7.0	864	1	Y63M_YEAST
19	78	6.9	215	1	NO22_SOYBN
20	77	6.8	523	1	HEX_ADEMI
21	77	6.8	909	1	CP5C_CANNA
22	76.5	6.8	393	1	IDH_STRMU
23	76.5	6.8	393	1	IDH_STRMU
24	76.5	6.8	724	1	DDX4_HUMAN
25	76	6.8	260	1	DCK_MOUSE
26	76	6.8	1031	1	YD69_SCHPO
27	76	6.8	1066	1	YD69_SCHPO
28	76	6.8	1066	1	YD69_SCHPO
29	75.5	6.7	1489	1	RNG2_SCHPO
30	75	6.7	414	1	Y878_METUA
31	75	6.7	424	1	PORA_CAMJE
32	75	6.7	626	1	HTPG_BUCBP
33	75	6.7	647	1	HTPG_MYCTU

34	75	6.7	3951	1	VGF1_IBVB
35	74.5	6.6	400	1	GTR3_RABIT
36	74.5	6.6	492	1	Y6CE_ECOLI
37	74.5	6.6	726	1	NPT1_YEAST
38	74.5	6.6	890	1	LPN1_HUMAN
39	74.5	6.6	2768	1	THYG_HUMAN
40	74	6.6	406	1	IDH_SPHYA
41	74	6.6	854	1	TRIC_SULTO
42	74	6.6	912	1	TBP1_HAEIN
43	73.5	6.5	355	1	UBPC_HUMAN
44	73.5	6.5	553	1	GRP2_HUMAN
45	73.5	6.5	2896	1	HCYG_OCTDO

## ALIGNMENTS

RESULT 1	ID	HBP2_RHIAP	STANDARD:	PRT:	190 AA.
AC	077421				
DT	30-MAY-2000	(rel. 39, Created)			
DT	30-MAY-2000	(rel. 39, Last sequence update)			
DT	28-FEB-2003	(rel. 41, Last annotation update)			
DE	Female-specific histamine-binding protein 2 precursor (FS-HBP2).				
OS	Rhhipcephalus appendiculatus (Brown ear tick).				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OX	Parasitiformes; Ixodidae; Ixodidae; Rhhipcephalus.				
RN	NCBI_TaxID=346311;				
RP	[1]				
RC	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).				
RX	TISSUE=Salivary gland;				
RA	MEDLINE=99288454; PubMed=10360182;				
RT	Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;				
RL	"Tick histamine-binding proteins: Isolation, cloning, and three-				
RL	dimensional structure.";				
CC	Mol. Cell 3:661-671(1999).				
CC	-1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO				
CC	OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO				
CC	SUPPRESS INFLAMMATION DURING BLOOD FEEDING.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	EMBL: U96081; AAC63107.1; -				
CC	PDB: 1QFT; 19-APR-00.				
CC	PDB: 1QFV; 19-APR-00.				
DR	InterPro: IPR002970; His_binding.				
DR	Pfam: PF02098; His_binding; 1.				
DR	ProDom: PD152455; His_binding; 1.				
KW	Signal; 3D-structure.				
FT	CHAIN	1	190		
FT	DISULFID	67	188		
FT	DISULFID	138	167		
FT	TURN	23	24		
FT	HELI	27	30		
FT	HELI	31	33		
FT	HELI	36	41		
FT	TURN	42	45		
FT	STRAND	48	53		
FT	STRAND	57	58		
FT	TURN	59	61		
FT	STRAND	62	63		
FT	TURN	64	64		
FT	STRAND	66	76		

2. FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN

Mon Aug 4 17:50:32 2003

dret6.rpr

Page 6

```

0Y 0EIERKEDDYVTSVFETFRNASSPIKYYNVETVAVAPGYGKNIRNALIEYQGGSLNT 140
Db 210 -----KGLMALITAAWROKHAANKVLRPEASTSGCVCPNPLDEIRLAGEYEYGVQNNNA 264
0Y 141 DTLIETDELCDFVFV-----PNADQGE 164
Db 265 TLIVDPDCKGLDSHQKTAFFALLVGIYLHALYKAKDDGATLPLPSVDAMLADNRDIG-E 323
0Y 165 LMYKSKSHYKHV 175
Db 324 LMMEKATYGHV 334

```

Search completed: August 4, 2003, 17:15:58  
Job time : 42 secs



A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yomi

Query Match 7.4%; Score 83; DB 2; Length 2285;  
Best Local Similarity 23.3%; Pred. No. 1.1e+02;  
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

OY 6 VLLTFVSALATQAETTSKAGENPLMAHEELLGKODAKSID--QGVSVYVIKAT 63  
DB 743 LVTSTLVGAFALG-----MALESLSFAEAKKAKDDEGSGQTINERIT 790  
OY 64 YENDTGSWSGQFCLOAEIERKEEDYTVTSFTFRNASSPIKYVNTETKAVFO--Y 120  
DB 791 NKDSTDKLIQOYK-----ELQVKESRSLTS-----DEGEYLAQVTOQLAQTFPALVK 838  
OY 121 GY-----KNRNAIEYOVGGGLNTDTLFTDGLCDVFPVNPADGCELWVK 169  
DB 839 GYDSQGNALIKTKNELEKALE-----NTRKEYLALKKQETRDŠAKKTFEDASKE--IKK 889  
OY 170 S-----HYKHVPDYCTFVFVNFCAKDRKTYDIFNEECVY 203  
DB 890 SKDELKQYKQIADYND-----KGRPKMDLIADDDY 920

## RESULT 9

C69192  
stomatol-like protein - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1999  
C:Accession: C69192  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Alredge, T.;  
Qiu, D.; Spedaleora, R.; Viciore, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jilwan, N.;  
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MWID:98037314; PMID:93711463  
A:Accession: C69192  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-318 <MTH>  
A:Cross-references: GB:AE000848; GB:AE000666; NID:g2621761; PIDN:AA85197.1; PID:g262177  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH692  
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 7.3%; Score 82.5; DB 2; Length 318;  
Best Local Similarity 23.3%; Pred. No. 9.2;  
Matches 38; Conservative 23; Mismatches 77; Indels 25; Gaps 6;

OY 5 VLLTFVSALATQAETTSKAGENPLMAHEELLGKODAKSIDGVSVTYVLA 64  
DB 10 VLLVLAFLKSLKIRPEKGV-----ERLGRYQ---RTVEGLVITIPFLIAIK 55  
OY 65 ENDTGSWSGQFCLOAEIERKEEDYTVTSFTFRNASSPIK-YVNTETKAVFOYGYK 123  
DB 56 KVP---MREQVVPVPPDEVITKDNVTVVVDCVIFEEVVDPFNNAVNVNVDYQAITKLAQT 112  
OY 124 NRNAIEYOVGGGLNTDTLFTDGLCDVFPVNPADGCELW 166  
DB 113 NLKNII-----GDELDQTL--TSREMINIQRLREVLEADATDKW 148

## RESULT 10

C97308  
probable acetyltransferase [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: C97308  
R:Kolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MWID:21359325; PMID:21359325

A:Accession: C97308

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-163 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81254.1; PID:g15026402; GSPDB:GND0168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:  
A:Gene: CAC3322

Query Match 7.3%; Score 82; DB 2; Length 163;  
Best Local Similarity 24.5%; Pred. No. 4.2;  
Matches 37; Conservative 11; Mismatches 51; Indels 52; Gaps 6;

OY 82 EIERKEEDYTVTSFTFRNASSPIKYVNTETKAV-----FOYGVKNI 125  
DB 3 KISVKQMDYETAKQITKWNTEKPSIYNMSECEIKELSLGYPFSAYDDKNYIVGYCF 62  
OY 126 RNAIEYOVGGGL-----NITDTLFTDGLCDVFPVNPADG-----CELWVKSHY 172  
DB 63 GEAQIIPVGQIGIYSYKNITDIGILNPLC-----GGGLGRDFCCGL----- 107  
OY 173 KHVPDYCTFVFVNFCAKD-RKTYDIFNEECV 202  
DB 108 -----DFARNTLCAKDFRLVATFNKPAI 131

## RESULT 11

T11796  
hypothetical protein A494R - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T11796  
R:Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z18606  
A:Accession: T11796  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-360 <GRA>  
A:Cross-references: EMBL:U42580; NID:g4028886; PIDN:ANC96861.1  
A:Experimental source: specific host Chlorella strain NC64A  
C:Genetics:  
A:Note: A494R

Query Match 7.3%; Score 82; DB 2; Length 360;  
Best Local Similarity 25.8%; Pred. No. 12;  
Matches 41; Conservative 19; Mismatches 59; Indels 40; Gaps 8;

OY 35 HEELGK---YODAMKSIDGVSVTYVLAATYENDTGSWSGQFCLOAEIERKEEDY 91  
DB 43 HEAVNINIQFOGKKKLENGV-----TKKTE-----KIQIEDEINYL 84  
OY 92 VTSFTFRNASSPIKYVNTETV-----KAVFOYGVKNIR-NAIE---YOVGGGLNT 140  
DB 85 LDTI-----PFIEYDVKETVSDVTEONSVEOVKSKTHNTFRKYLFEHVEKVSNP 136  
OY 141 DTLFTDGLCDVFPVNPADGCELWVKSHYKHVPDYC 179  
DB 137 TLDAVTDREIVDQITCTCGGOMELWVNSIQSDLVNCEC 175

## RESULT 12

T18876  
hypothetical protein C03C10.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2000  
C:Accession: T18876  
R:Barks, M.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19036  
A:Accession: T18876





A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.  
A:Reference number: A70100; MUID:98065943; PMID:9403685  
A:Accession: G70169  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-650 <KLE>  
A:Cross-references: GB:AE001157; GB:AE000783; NID:g268847L; PID:g2688447  
A:Experimental source: strain B31  
R:Ojalmt, C.; Davidson, B.E.; Saint Gilrons, I.; Old, I.G.  
Microbiology 140, 2931-2940, 1994  
A:Title: Conservation of gene arrangement and an unusual organization of rRNA genes in the  
A:Reference number: I40241; MUID:95111614; PMID:7812434  
A:Accession: I40247  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 531-609 <RES>  
A:Cross-references: GB:E32145; NID:g476004; PID:AAAC4403.1; PID:g476594  
C:Genetics:  
A:Gene: htpG  
C:Superfamily: heat shock protein 90

Query Match	8.5%	Score 96	DB 2	Length 650
Best Local Similarity	21.5%	Pred. NO	1.4	
Matches 46	Conservative 34	Mismatches 92	Indels 42	Gaps 8

0y 11FVSAALATQ-AETTSKAGENP--LMAHELLGKODANSIDGVSVTVYLAITYEND 67  
| | : : : | - - - : : : |  
Db 155 FYSAFIVSEKVEVTSKALESDAIVWSOQTGEIEKAKKEESCTEIKILYNLEGLE-- 212

```

0Y 68 TSGMSGSGFCLOV-----QEIEREEDYVTSVETFRNA 101
      :::::P:::
213 ---YANKMKIQEIIKKYSNHINIPYIKYSEPIIMKDGKQECIEEKEKNETALTMTKN- 268

```

QY 102 SSPKRYNNTETVFAVFGYGRKKNRAIEYQVGGNLINDTLFTDGEICDVFVYPNADQ 161  
 269 KSEIKAEENEFYKNT-FTDYENLMHINTKAECNLLEYRNLFYPPSKAPYDLYI-PTNPK 326

QY	162	GCELMVKK-----SHYKHAVDPDYCFEVFNNFCARD	190
		: : : :   : : : :	
Db	327	GVKLEINRIFITDSEGSLLPNYLRFRIKGIIDCD	360

### RESULT 3

CDate: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
CAccession: F71258  
CSpecies: *Treponeuma pallidum* (synth subs. *pallidum* spirochete)  
CDate: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
CAccession: F71258  
BAccession: C.M.: Morris S.J.: Weinstein G.M.: White O.: Sutton G.G.: Dodson R.

the snail's ear cochlea  
Article: Complete genome sequence of *Trematode pallidum*  
Science 281, 375-388, 1998  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
reon, J.; Kralak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback  
Article: Complete genome sequence of *Trematode pallidum*  
Science 281, 375-388, 1998  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

A:Accession: F71258  
A:Reference number: A71250; MUID:9833270; PMID:9665876  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-693 <COL>  
A:Cross-references: GB:AE001265; NID:g3323295; PIDN:AAC65938.1; P  
A:Experimental source: strain Nichols  
C:Genetics:

Query Match      7.7%      Score 87      DB 2:      length 639:

[illegible]

Db	182	VVTLHLSOENSEFATR-----	VRLEEVYAKYSDHIAPDIYLHLQKEYDKDGA	228
QY	55	VTVVLAKTTYENDTGSWCSQFCKLOVEIERKEEDY----	TVTSVETFRNASSPIDIKYNN	109

Db	229	VTDTQKKVDPVDNAGALMKRPKS-----ELKEEDYHRRFYQTLT-----RSTSPRLLYVH	277
Qy	110	VTVETVAKAVROYGYKINIRNAIEYQVGSGNLITDTLFTDGG--ELCDVFPV-----NA	155
Db	278	TK-----AGTQDEYVTLFVYPAKAPDDELFHA	303
Qy	160	D--QGCELMYKK-----SHYKAVPDPXCFEVNVFCAKD	190
Db	304	DYKPGVKLFYKRVFTIDDEKELLPLYLRKVRGVDSIED	341

RESULT 4  
BA9132  
fat facets (faf) splice form 1 - fruit fly (*Drosophila melanogaster*)

CiDate: 19-Dec-1993 #sequence\_revision 25-Apr-1997 #text\_change 01-Dec-2000  
CiAccession: B49132; A49132  
RiFischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.

A:Title: The fat facets gene is required for Drosophila eye and embryo development  
A:Reference number: A49132; MUID:93202020; PMID:1295747  
A:Contents: Isogenic st

A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-2747 <FIS>

A1Note: sequence inconsistent with the nucleotide translation  
A1Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCEP:129008)  
A1Accession: A49132

A:Molecule type: nucleic acid  
A:Residues: 1-2704, 'VT', 2707, 'ANNV' <FI2>  
A:Cross-references: GB:I04958; NID:g157410; PIDN:AAF01346.1; PID:56013475  
A:Notes: Sequences extracted from yeast poly(A) library (MIM:137025)  
MIM:137025

C:Keywords: alternative splicing

Matches	54;	Conservative	28;	Mismatches	72;	Indels	82;	Gaps	12
Oy	6	VLLTFTVSAALATQAEFTTSKAGENDLMAHBEELG	----	KYDA	-----	WKSTOG	52		

Db 1756 VWLKHVQAFHHLGHSALQYVPRSLMTHKRLGLGEVNLRRQGDVAEEFMSLLESDG 181

Qy 53 VSVIYVLAKTYENDT--GSGSGQPKCLOVEIERKEEDYIYTSVFTRNASSPIKYYNV 110

Db 1816 LK---ALGQQLMNTLTGSGSDQKLCQECCHRYSKPEFVSFV-DIRNHSS-----L 1866

Qy 111 TETVKAVFQYGYKKNIRNAIEQVGGGINTPTLIFTDGEL---CDVFYVPNADQGCCLAV 1675

```

Db      1866 TESLEQ-----LYVGELEEGADAHYCDKCDKKV-VTV 1899
Qy      168 KKSRYKHP-----DY--CTFENVFCADRKRTYDIFNEECVYNGEPM 208

```

DB 1897 KRVCKVKKLPPVLAIDJLKRFEQYERVCALIKEN-----DYFEEPRILDMEPY 1942

AB7112  
heat shock protein Hsp90 family [Imported] - *Mycobacterium leprae*  
C:Species: *Mycobacterium leprae*  
C:Date: 20-Apr-2001 #sequence revision 20-Apr-2001 #text change 10-May-2001

C:accession: AB7112  
R:cole, S.T.; Eglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler  
R.; Davies, R.M.; Devlin, K.; Dutnoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.  
eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squan  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86809: MUID:21128732: PMID:11234002

A: Accession: A8/112

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 17:13:11 ; Search time 40 Seconds

(without alignments)  
502.481 Million cell updates/sec

Title: DRET6

Perfect score: 1125

Sequence: 1 MKMOVLLTFVSAALATQA.....DRKTYDIFNECVNGEPWL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_76:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96.5	8.6	621	2	heat shock protein
2	96	8.5	650	2	heat shock protein
3	87	7.7	639	2	probable heat shock
4	86	7.6	2747	2	fat facets (fat) s
5	85	7.6	656	2	heat shock protein
6	84.5	7.5	621	2	heat shock protein
7	84.5	7.5	917	2	autolysin, amidase
8	83	7.4	2285	2	probable transglyc
9	82.5	7.3	318	2	stomatin-like prot
10	82	7.3	163	2	probable acetyltra
11	82	7.3	360	2	hypothetical prote
12	82	7.3	381	2	hypothetical prote
13	82	7.3	390	2	conserved hypotet
14	82	7.3	637	2	trig protein - Ent
15	82	7.3	637	2	trig protein - Esc
16	81.5	7.2	627	2	90kDa chaperone -
17	80.5	7.2	621	2	chaperone and heat
18	79.5	7.1	621	2	hypothetical prote
19	79.5	7.1	725	2	hypothetical prote
20	78.5	7.0	455	2	hypothetical prote
21	78.5	7.0	864	2	hypothetical prote
22	78	6.9	215	2	nodulin-22 precurs
23	78	6.9	523	1	cytochrome P450 52
24	78	6.9	1871	2	hypothetical prote
25	77.5	6.9	223	2	ribonuclease (EC 3
26	77.5	6.9	506	2	hypothetical prote
27	77	6.8	241	2	hypothetical prote
28	77	6.8	331	2	probable periplasm
29	77	6.8	780	2	probable heat shoc

30	77	6.8	914	2	transferrin-binding
31	76.5	6.8	587	2	poly-beta-hydroxyb
32	76.5	6.8	635	2	probable RNA helic
33	76.5	6.8	807	2	hypothetical prote
34	76.5	6.8	2819	2	conserved hypotet
35	76	6.8	260	2	deoxycytidine kina
36	76	6.8	510	2	hypothetical prote
37	76	6.8	510	2	lipase-like protei
38	76	6.8	531	2	hypothetical prote
39	76	6.8	598	2	hypothetical prote
40	76	6.8	781	2	probable arylsulfa
41	76	6.8	1031	2	heat-shock protein
42	76	6.8	1066	2	probable GTPase ac
43	76	6.8	1068	2	hyaluronidase [imp
44	75	6.8	1277	2	hypothetical prote
45	75.5	6.7	820	2	hypothetical prote
					probable transmem

## ALIGNMENTS

## RESULT 1

E97862 heat shock protein htpg [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C:Accession: F97862

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;

Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: F97862

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-621 <KUR>

A:Cross-references: GB:AE006914; PIDN:AAL03840.1; PID:915620441; GSPDB:GN00173

A:Genetics:

A:Gene: htpg

C:Superfamily: heat shock protein 90

Query Match 8.6%; Score 96.5; DB 2; Length 621;

Best Local Similarity 21.0%; Pred. No. 1.2;

Matches 43; Conservative 46; Mismatches 81; Indels 35; Gaps 10;

QY	11	FVSALATQAETTSKAGENP--LMAHEILGKY--ODAMKSIOGVSVTYLAKTYYEN 66
DB	127	YSEFWADKVTYTSKAGEDVHTM-ESDGLGETVSDSKXETRTGTEIVLHIKK---EE 182
QY	67	DTGWSGQFKCLOVEIERKEEDYTVTSVFTF-----RNASSPIKYYNVTETKAV 117
DB	183	DT--FLDHF--LKHIVKSYSDHIAVPIFFDEAGNNEIDLSMSALMTPKSEITPEQ 237
QY	118	FQYGYKIRNAIE-----YVGGLNITPTLFTDGLCDVFRVPAADGCELYAKK- 169
DB	238	YKEFYKSLSYAIDDPWITMHNKNEGALIEFTMLFIPSSKTPDLPH-PDRKRRVRLYIKRV 296
QY	170	----SHYKVPDYCFVFNVCARD 190
DB	297	FISDENIDLPSYLFRLRGVDSDD 321

## RESULT 2

G70169 heat shock protein 90 (htpg) homolog - Lyme disease spirochete

N:Alternate names: C62.5 heat shock protein

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 20-Aug-1999

C:Accession: G70169; I40247

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh

son, D.; Peterson, J.; Kertlage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Yu

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 17:16:01 ; Search time 52 Seconds  
(without alignments)  
477.323 Million cell updates/sec

Title: DRET6  
Perfect score: 1125  
Sequence: 1 MKMGVLLTFVSLATATQA.....DRKTYDFNECCVNGEPPML 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA.\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	282	25.1	171	US-10-085-572-7	Sequence 7, Appl1
2	282	25.1	171	US-10-087-195-7	Sequence 7, Appl1
3	251.5	22.4	182	US-10-085-572-8	Sequence 8, Appl1
4	251.5	22.4	182	US-10-087-195-8	Sequence 8, Appl1
5	239	21.2	172	US-10-085-572-6	Sequence 6, Appl1
6	239	21.2	172	US-10-087-195-6	Sequence 6, Appl1
7	102.5	9.1	722	US-09-815-242-10796	Sequence 10796, A
8	97.5	8.7	221	US-09-728-914-4	Sequence 4, Appl1
9	84	7.5	217	US-09-728-914-22	Sequence 22, Appl1
10	83	7.4	2285	US-09-932-183A-2	Sequence 2, Appl1
11	81	7.2	659	US-10-090-624-12	Sequence 12, Appl1
12	79.5	7.1	913	US-10-043-344-5	Sequence 5, Appl1
13	77	6.8	914	US-10-043-344-11	Sequence 11, Appl1
14	75	6.7	912	US-10-043-344-7	Sequence 7, Appl1
15	75	6.7	912	US-10-043-344-9	Sequence 9, Appl1

16	74.5	6.6	890	US-10-028-056-3	Sequence 3, Appl1
17	73	6.5	1417	US-09-753-143-78	Sequence 78, Appl1
18	72.5	6.4	911	US-10-043-344-107	Sequence 107, App
19	72.5	6.4	2756	US-10-331-061-7	Sequence 7, Appl1
20	72	6.4	507	US-09-738-626-5458	Sequence 5458, Ap
21	72	6.4	956	US-10-121-032-63	Sequence 63, Appl1
22	72	6.4	956	US-10-093-037-63	Sequence 63, Appl1
23	71.5	6.4	564	US-10-081-872-120	Sequence 120, App
24	71.5	6.4	891	US-10-028-056-4	Sequence 4, Appl1
25	71.5	6.4	924	US-10-028-056-5	Sequence 5, Appl1
26	71.5	6.4	2012	US-09-808-602-68	Sequence 68, Appl1
27	71.5	6.4	2012	US-09-800-198-57	Sequence 57, Appl1
28	71	6.3	550	US-10-081-872-106	Sequence 106, App
29	71	6.3	618	US-10-295-403-166	Sequence 166, App
30	71	6.3	619	US-09-934-455-142	Sequence 142, App
31	71	6.3	677	US-09-862-027-38	Sequence 38, Appl1
32	70.5	6.3	451	US-09-759-130B-413	Sequence 413, App
33	70.5	6.3	451	US-10-042-431-43	Sequence 43, Appl1
34	70.5	6.3	470	US-09-759-130B-439	Sequence 439, App
35	70.5	6.3	470	US-10-042-431-69	Sequence 69, Appl1
36	70	6.2	247	US-09-738-626-3674	Sequence 3674, Ap
37	70	6.2	576	US-09-731-175-4	Sequence 4, Appl1
38	70	6.2	576	US-10-011-348-30	Sequence 30, Appl1
39	70	6.2	638	US-09-842-758-22	Sequence 22, Appl1
40	70	6.2	853	US-10-156-761-8930	Sequence 8930, Ap
41	69.5	6.2	225	US-09-847-208-110	Sequence 110, App
42	69.5	6.2	373	US-09-759-130B-376	Sequence 376, App
43	69.5	6.2	373	US-10-042-431-6	Sequence 6, Appl1
44	69.5	6.2	374	US-10-270-875-39	Sequence 39, Appl1
45	69.5	6.2	374	US-10-270-878-39	Sequence 39, Appl1

## ALIGNMENTS

RESULT 1				
US-10-085-572-7				
; Publication 7, Application US/10085572				
; Sequence No. US20020151499A1				
GENERAL INFORMATION:				
; APPLICANT: Nuttall, Patricia, Ann				
; APPLICANT: Paesen, Guido, Christiaan				
; TITLE OF INVENTION: Treatment of Conjunctivitis				
; FILE REFERENCE: 2488-1-003				
; CURRENT APPLICATION NUMBER: US/10/085,572				
; CURRENT FILING DATE: 2002-02-27				
; PRIOR APPLICATION NUMBER: PCT/GB00/03282				
; PRIOR FILING DATE: 2000-08-24				
; PRIOR APPLICATION NUMBER: 9920674.0				
; PRIOR FILING DATE: 1999-09-01				
; NUMBER OF SEQ ID NOS: 8				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 7				
; LENGTH: 171				
; TYPE: PRT				
; ORGANISM: FS-HBP 2				
US-10-085-572-7				
Query Match				
Best Local Similarity 37.5%; Pred. No. 7.3e-23;				
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;				
QY	31	PLMAHEELGKQDADKSIDGVSVTYVLAQTYENDTSGMSQSFCLQVOELERKEEY 90		
DB	3	PMDADPANGAHODAKSLKADVENYVYKATYKNDP-VMGNDPFCVGMANDVDEDE 61		
QY	91	TVTSVTF-RNASSPRTKYVNTETVKAFOYGRKRNMAIEYOVGGALITDPLTFDDE 149		
DB	62	SIQAEELFMNNDTNNQF--ATEKYAVAMKYG-NENMAFRYETEDGQFTDVIASD-D 117		
QY	150	LCDFVFPVPAAD--QCCELAVKSKYKHVPDCTEFPVNFCAKDRKTYDIFNECV 202		
DB	118	NCDFVFPVPAAD--QCCELAVKSKYKHVPDCTEFPVNFCAKDRKTYDIFNECV 202		

## RESULT 2

US-10-087-195-7  
; Sequence 7, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: FS-HBP 2  
US-10-087-195-7

Query Match 25.1%; Score 282; DB 14; Length 171;

Best Local Similarity 37.5%; Pred. No. 7, 3e-23;  
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;

QY 31 PLMAHELLGKYDAMKSIDQVSVTVYLAKTTEYNDTSGMSQFCKLOV--QETERKEEDY 90  
DB 3 PDMADEANGAHODAMKSLKADVENYVYKATYKNDP--VWGNDFCTCVGMADVNEDEK 61  
QY 91 TVTWSVTFE--RNASSPIKYYVETVKAPOYGYKKNIRNALEYOVGGGLNTDILIFDGE 149  
DB 62 SDAEPLFNNAADTNQOF--ATEKVTAVMKYG--NRENAERYETEDQVETDVIAVSD-D 117  
QY 150 LCDVFPVFNAD--OGCELVWKKSHYKHPDYCTFEVFNVCADKRTYDIFNECY 202  
DB 118 NCVIIVYVPGTGDGNEGEELMT--TDYDNIPANCLNPFNE--AVGRETRDVFISACL 170

## RESULT 3

US-10-085-572-8  
; Sequence 8, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; APPLICANT: Paesen, Guido, Christiaan  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: MS-HBP 1  
US-10-085-572-8

Query Match 22.4%; Score 251.5; DB 14; Length 182;  
Best Local Similarity 35.3%; Pred. No. 1, 7e-19;  
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

QY 30 NPLMAHELLGKYDAMKSIDQVSVTVYLAKTTEYNDTSGMSQFCKLOV--QETERKE 87  
DB 1 NPTWANEARKLSGYDAMKSLQDQNKRYLAQATQTTD--GVWGEFTCVSVTAEKIGKK 59

QY 88 EDYTVTSVTFERNAS--SPIKYVNVETVKAFOYGYKKNIRNALEYOVGGGLNT--DTLI 144

DB 60 LNATI-----LYKKNHLTDLKEHSHETITVWKAIDY---TTENGIKYETQGRTRQTFEDVEV 112

QY 145 FTDEGLCDVFPV--NADQG--CELVWKKSHYKHPDYCTFEVFNVCADKRT 193

DB 113 FSDYKNCDVLFVPERKSGSDGDYELWVSEDKIDKIPDCKFTMAVFAQOQDEK 165

## RESULT 4

US-10-087-195-8  
; Sequence 8, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; APPLICANT: Paesen, Guido, Christiaan  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: MS-HBP 1  
US-10-087-195-8

Query Match 22.4%; Score 251.5; DB 14; Length 182;

Best Local Similarity 35.3%; Pred. No. 1, 7e-19;  
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

QY 30 NPLMAHELLGKYDAMKSIDQVSVTVYLAKTTEYNDTSGMSQFCKLOV--QETERKE 87  
DB 1 NPTWANEARKLSGYDAMKSLQDQNKRYLAQATQTTD--GVWGEFTCVSVTAEKIGKK 59  
QY 88 EDYTVTSVTFERNAS--SPIKYVNVETVKAFOYGYKKNIRNALEYOVGGGLNT--DTLI 144  
DB 60 LNATI-----LYKKNHLTDLKEHSHETITVWKAIDY---TTENGIKYETQGRTRQTFEDVEV 112  
QY 145 FTDEGLCDVFPV--NADQG--CELVWKKSHYKHPDYCTFEVFNVCADKRT 193  
DB 113 FSDYKNCDVLFVPERKSGSDGDYELWVSEDKIDKIPDCKFTMAVFAQOQDEK 165

## RESULT 5

US-10-085-572-6  
; Sequence 6, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; APPLICANT: Paesen, Guido, Christiaan  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 172  
; TYPE: PRT  
; ORGANISM: FS-HBP1  
US-10-085-572-6

Query Match 21.2%; Score 239; DB 14; Length 172;

	Best Local Similarity	34.5%, Pred. No. 3.7e-18;	Matches 61;	Conservative 24;	Mismatches 82;	Indels 10;	Gaps 7;
Qy	29 ENPLMAHELGGKTYODAMKSIDQGVSYTYVLAKTTYENDTGSWGSQFCLQVOETERKEE	88					
Db	1 DKPWAEBAANGEHODAMKHLQKLVENYDILKATYKNDP--VWGNDFTCVGTAQNLNED	59					
Qy	89 DYTTSVTEFPNASSPIKYVWTEYVKAVFQYGYKNINAILEYOGGGINTDILTFIDG	148					
Db	60 EKNVEMFMFNNAADTV-YQHTFEKATPDKKMIGY-NKENALITYQTEDQGVLTDLVAFSD-	116					
Qy	149 ELCDVFVYPNAD---OGCELWVKRSHYKHPDYCTFEVNFVCAMDKRKYTDIFNEBCV	202					
Db	117 DNGCYVYVALGPDGSGAGYELMA--TDYTDVDPASCLEKKNERYAA-GLPYRDVYTSCL	170					
RESULT 6							
US-10-087-195-6							
Sequence 6, Application US/10087195							
Publication No. US20020193306A1							
GENERAL INFORMATION:							
APPLICANT: Nuttall, Patricia, Anne							
APPLICANT: Praesen, Guido, Christian							
TITLE OF INVENTION: Treatment of Allergic Rhinitis							
FILE REFERENCE: 2488-1-004							
CURRENT APPLICATION NUMBER: US/10/087, 195							
CURRENT FILING DATE: 2002-03-01							
PRIOR APPLICATION NUMBER: PCY/GB00/03287							
PRIOR FILING DATE: 2000-08-24							
PRIOR APPLICATION NUMBER: 9920673.2							
PRIOR FILING DATE: 1999-09-01							
NUMBER OF SEQ ID NOS: 8							
SOFTWARE: FastSeq for Windows Version 4.0							
SEQ ID NO 6							
LENGTH: 172							
TYPE: PRF							
ORGANISM: PS-HBPI							
US-10-087-195-6							
Query Match	21.2%, Score 239; DB 14; Length 172;						
Best Local Similarity	34.5%, Pred. No. 3.7e-18;						
Matches 61; Conservative 24; Mismatches 82; Indels 10; Gaps 7							
Qy	29 ENPLMAHELGGKTYODAMKSIDQGVSYTYVLAKTTYENDTGSWGSQFCLQVOETERKEE	88					
Db	1 DKPWAEBAANGEHODAMKHLQKLVENYDILKATYKNDP--VWGNDFTCVGTAQNLNED	59					
Qy	89 DYTTSVTEFPNASSPIKYVWTEYVKAVFQYGYKNINAILEYOGGGINTDILTFIDG	148					
Db	60 EKNVEMFMFNNAADTV-YQHTFEKATPDKKMIGY-NKENALITYQTEDQGVLTDLVAFSD-	116					
Qy	149 ELCDVFVYPNAD---OGCELWVKRSHYKHPDYCTFEVNFVCAMDKRKYTDIFNEBCV	202					
Db	117 DNGCYVYVALGPDGSGAGYELMA--TDYTDVDPASCLEKKNERYAA-GLPYRDVYTSCL	170					
RESULT 7							
US-09-815-242-10796							
Sequence 10796, Application US/09815242							
Patent No. US20020061569A1							
GENERAL INFORMATION:							
APPLICANT: Haselbeck, Robert							
APPLICANT: Ohlsen, Karl L.							
APPLICANT: Zyskind, Judith W.							
APPLICANT: Wall, Daniel							
APPLICANT: Trawick, John D.							
APPLICANT: Carr, Grant J.							
APPLICANT: Yamamoto, Robert T.							
APPLICANT: Xu, H. Howard							
TITLE OF INVENTION: Identification of Essential Genes in							
TITLE OF INVENTION: Prokaryotes							
FILE REFERENCE: ELITRA.011A							
CURRENT APPLICATION NUMBER: US/09/815.242							

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for windows version 4.0
; SEQ ID NO: 10796
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10796

Query Match          9.1%; Score 102.5; DB 9; Length 722;
Best Local Similarity 26.9%; Pred. No. 0.023;
Matches 60; Conservative 26; Mismatches 72; Indels 65; Gaps 15;

QY 16 LATQAEITSA-----KAENELPMABEELIKYQDAWKSI-DQGVSVTVVLAKTTY----- 64
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 204 LVGDDETSSSGFAVKKQNP----ELIKKFNAGLKLKLDG--TYDKLNNLATGD 255
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 ENDIGSMGSPKCLQVOEIERKEEDYTVTSVTFPMNASPIKYYNV-----TEYKA 116
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 256 ETNQDAEQK-----KIKPKKEKYYIASDSTF-----APEFQNAQGDYGVIDVLYKR 306
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 117 V-----FOY-GYKIRNAIEY-QVGC---GLNTIDTLITFDGELCDVFFYNADQG 162
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 307 AAELQGFTEVERKEFIFSSAYOVAVESGADGMVAGMTITD----DRKRAFDFSVYFDSG 361
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 163 CELMWKSKH--YKHVPDYCTFEFVN-----FCAKDRKTYD 195
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 362 IQIAVKKGNDKIKSYDDLKGGKVKYKIGTESADFLKKKKKYD 404
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-09-728-914-4
; Sequence 4, Application US/09728914
; Patent No. US20010046499A1
; GENERAL INFORMATION:
; APPLICANT: KANTOR, FRED S.
; APPLICANT: KANTOR, FRED S.
; APPLICANT: FIKRIG, EROL
; APPLICANT: DAS, SUBRATA
; TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
; FILE REFERENCE: YU-107
; CURRENT APPLICATION NUMBER: US/09/728,914
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,048
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/240,716
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Ixodes scapularis
US-09-728-914-4

Query Match          8.7%; Score 97.5; DB 9; Length 221;
Best Local Similarity 23.4%; Pred. No. 0.015;
Matches 49; Conservative 43; Mismatches 80; Indels 37; Gaps 14;

```

```

OY      10 TVFAAALATQAE-----TSAGKNPILWAHELGLKY--ODAMSIDOGSVFVLA 60
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      31 TYSGTGTTTRRGTCGARNAITVTAPBEDP-----SKTEONATRYVE--MATAOWK 80
OY      61 KTYENDTGSWSQFKLOAELEKEEDYTVTSV-PTFNASSPIKYWTFET--VKAV 117
       || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      81 WRITYD-VTFDSGNINVOQENFRVMEKR---TPNYSQRYRKSKN-SWEIIDEFLIKDI 134
OY      118 FQYGKINRAIEYQVGCGNLITDTLL-FPDGELCDVFYYPNADQC---CELWKK-SHY 172
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      135 GEHGFPMVMNEQRPPIG---IATDNLYASNYVCVLRIPTNOGERHCIDLMANLITS 191
OY      173 KHVPDYCFEYNFCADKDRKYDJFNEEC 201
DB      192 QEPDDCLNKFEFYC-NFTQIRVRYPSQ 219

RESULT 9
US-09-728-914-22
; Sequence 22, Application US/09728914
; Patent No. US20010046499A1
; GENERAL INFORMATION:
; APPLICANT: KANTOR, FRED S.
; APPLICANT: FIRKIG, EROL
; APPLICANT: DAS, SUBRATA
; TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
; TITLE OF INVENTION: THEM
; FILE REFERENCE: YU-107
; CURRENT APPLICATION NUMBER: US/09/728, 914
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169, 048
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/240, 716
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 22
LENGTH: 217
TYPE: PRP
ORGANISM: Ixodes scapularis
US-09-728-914-22

Query Match          7.5%; Score 84; DE 9; Length 217;
Best Local Similarity 24.1%; Pred. No. 0.45;
Matches 27; Conservative 17; Mismatches 48; Indels 20; Gaps 3;

OY      99 RNASSPIKYNVETPKAVFOYGKINRNAIEYOVGGCLNTDTLLFTOS----- 148
       || || | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      79 RNDSS-----YEFVKRTLSTMSVGEQSPSRHTIEILTSGSKSEKVTKYEYEDRNTRY 134
OY      149 --ELCDVEFYYPNAD---QGCELVKKSHYKHPDYCTEFVNFCAKDRTTY 194
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      135 ADTRCLILRISRTQKVPLRSCDLLMWKKTFLKNPLRHCRRLFDFVFCMMRRDEF 186

RESULT 10
US-09-932-183A-2
; Sequence 2, Application US/09932183A
; Patent No. US20020127641A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394C1-US
; CURRENT APPLICATION NUMBER: US/09/932,183A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/308, 375
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US98/18828
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: EP9719636.4
; PRIOR FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
```

[illegible]









Db 1 KLAIAVCVFISTVSSVPSDNTSEE-----EDLNTKRDIMKTSK--SKVFWLRL 49  
QY 62 TTYENDTSGWSQFCL-----QVDEI-----ERKEDYTVTVF--TFRNASSP 104  
Db 50 RTWVD-----GKRSCLRGTGLKRDKNHRIEIMSYDSQETLYTKYITRLGRSG 105  
QY 105 IKYVNTFVKAVFOYGYKNIRNAIEYOVGGGLNTDTLFTDGLCDVFFV----- 156  
Db 106 DRNH-----MGVSLQGYNH--TGIEYK-----MYDDQCALIKYTKDNRNQP 147  
QY 157 PNADGCELWKKSHKHVPDY--CTFEVNFPCAKDRKTYD 195  
Db 148 QNLKACEMWATANDANSVNSIACEVYVORCNPNNNSVD 188

## RESULT 2

US-09-107-532A-5279  
; Sequence 5279, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arinello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5279:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 497 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (b) LOCATION 1...497  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5279:  
US-09-107-532A-5279  
Query Match 7.5%; Score 84.5; DB 4; Length 497;  
Best Local Similarity 22.6%; Pred. No. 1.2;  
Matches 47; Conservative 29; Mismatches 85; Indels 47; Gaps 9;  
QY 26 KAGENPLMAHEELGKYODAMKSINOVSIVTYVLAITYEND--TGSMSQKRCLOVDI 83  
Db 15 KPEBGLMGATPANOCEGAMVDVDEKGLSVSDY--TFDSLPKPKKTDQHMHTHOQV 71

QY 84 ERKEDYTVTVTFEFNASSPIKYVNTETVKAIVFOGYKNIRNAIEYOVGGGLNTDTL 143  
Db 72 KANQP-----NSKYYKRRHNDYHHFKEDIRLFAIMGFCRYMSIAM-----TR 118  
QY 144 IF-----TDGLCDVFFVPNDQ-----GCELWKKSHY-----KHVPD 177  
Db 119 IFPHGDEETPNEAGLFYDQVFDECLKYGIEPVLSIHEMPLYLVTEGGMFRRLIOF 178  
QY 178 YCTFEVNFPCAKDRKT--YDIENE-ECV 202  
Db 179 YRFAETVFRKRYKNKYWMTFENEINCV 206

## RESULT 3

US-09-308-375-2  
; Sequence 2, Application US/09308375  
; Patent No. 6300117  
; GENERAL INFORMATION:  
; APPLICANT: Genencor International, Inc.  
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms  
; FILE REFERENCE: GC394-PCT  
; CURRENT APPLICATION NUMBER: US/09/308,375  
; CURRENT FILING DATE: 1999-05-14  
; EARLIER APPLICATION NUMBER: EP9719636.4  
; EARLIER FILING DATE: 1997-09-15  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 2285  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
US-09-308-375-2

Query Match 7.4%; Score 83; DB 4; Length 2285;  
Best Local Similarity 23.3%; Pred. No. 16;  
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

QY 6 VLLTFVSAALATQAEETSAKAGENPLMAHEELGKYODAMKSID--QGVSYVYVLAITY 63  
Db 743 LVTSTLVGAFALG-----MALESISFAKAKKAKDFFESQQTINVEAITT 790  
QY 64 YENDTSGWSQFCLQVDEIERKEDYTVTVTFEFNASSPIKYVNTETVKAIVFO--Y 120  
Db 791 NKDSTDKLIQYK-----ELQVKESRSLTS-----DEQEYLVQTOQLQTFPALVK 838  
QY 121 GY-----KNIRNAIEYQVGGGLNTDTLFTDGLCDVFFVPNDQGCCELWKK 169  
Db 839 GYDSQGNALTKNKELEKAI-----NTKEYLALKQETRDSAKTFEDASKE--IRK 889  
QY 170 S-----HYKHVPDYCTFEVNFPCAKDRKTYDIFNEECVY 203  
Db 890 SKDELKQYKQIADYND-----KGRPKMDLIADDDY 920

## RESULT 4

US-08-894-818B-1  
; Sequence 1, Application US/08894818B  
; Patent No. 6261822  
; GENERAL INFORMATION:  
; APPLICANT: TAKAKURA, Hikaru  
; APPLICANT: MORISHITA, Mio  
; APPLICANT: YAMAMOTO, Katsuhiko  
; APPLICANT: MITTA, Masanori  
; APPLICANT: ASADA, Kiyoko  
; APPLICANT: TSUNASAWA, Susumu  
; APPLICANT: KATO, Ikuooshin  
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Browdy and Neimark  
; STREET: 419 Seventh Street N.W., Ste. 300  
; CITY: Washington

STATE: D.C.  
COUNTRY: United States of America  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,818B  
FILING DATE: 20-MAY-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03253  
FILING DATE: 07-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 323285/1995  
FILING DATE: 12-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Browdy, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: TAKAKURA-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 659 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-818B-1

Query Match 7.2%; Score 81; DB 3; Length 659;  
Best Local Similarity 25.9%; Pred. No. 4.2;  
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;

QY 41 KYDAMK-----SIDGVSV-TYVLAKTYENDTGSW--GSQKCLQVDEIERKEDYT 91  
DB 436 KYDYARALFTGSVADGASATFPDVSQATFVATLTYWDGSSDIDLTYLDPNGNEVDYS 495  
QY 92 VTSVETFRNASSPIKYYNVE---TVKAVFQYKKNIRNAIEYOV-----GGG- 136  
DB 496 YTAIYGFEEK---VGIYNPTAGTWTAKVY---STKGAAN---YQVDVYSDGSLQSGG 545  
QY 137 -----LINTDTLFTDGLCDVFPVNPADGCELMVKSHYKHVPDYCTFVFNYFC 187  
DB 546 PNPMPNPPTPTDTQFT--GSVND--YMDTSD-----FTTMNVNS 583  
QY 188 AKDRKTYDI-----FNEECYNGEP 207  
DB 584 GATKITGDLTFDTSYNDLDLYDP 608

RESULT 5  
US-08-894-818B-5  
Sequence 5, Application US/08894818B  
Patent No. 6261822  
GENERAL INFORMATION:  
APPLICANT: TAKAKURA, Hikaru  
APPLICANT: MORISHITA, Mio  
APPLICANT: YAMAMOTO, Katsuhiko  
APPLICANT: WITTA, Masanori  
APPLICANT: ASADA, Kiyozo  
APPLICANT: TSUNASAWA, Susumu  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Browdy and Nelmark  
STREET: 419 Seventh Street N.W., Ste. 300  
CITY: Washington

STATE: D.C.  
COUNTRY: United States of America  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,818B  
FILING DATE: 20-MAY-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03253  
FILING DATE: 07-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 323285/1995  
FILING DATE: 12-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Browdy, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: TAKAKURA-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 659 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-818B-5

Query Match 7.2%; Score 81; DB 3; Length 659;  
Best Local Similarity 25.9%; Pred. No. 4.2;  
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;

QY 41 KYDAMK-----SIDGVSV-TYVLAKTYENDTGSW--GSQKCLQVDEIERKEDYT 91  
DB 436 KYDYARALFTGSVADGASATFPDVSQATFVATLTYWDGSSDIDLTYLDPNGNEVDYS 495  
QY 92 VTSVETFRNASSPIKYYNVE---TVKAVFQYKKNIRNAIEYOV-----GGG- 136  
DB 496 YTAIYGFEEK---VGIYNPTAGTWTAKVY---STKGAAN---YQVDVYSDGSLQSGG 545  
QY 137 -----LINTDTLFTDGLCDVFPVNPADGCELMVKSHYKHVPDYCTFVFNYFC 187  
DB 546 PNPMPNPPTPTDTQFT--GSVND--YMDTSD-----FTTMNVNS 583  
QY 188 AKDRKTYDI-----FNEECYNGEP 207  
DB 584 GATKITGDLTFDTSYNDLDLYDP 608

RESULT 6  
US-09-445-472-12  
Sequence 12, Application US/09445472  
Patent No. 6358726  
GENERAL INFORMATION:  
APPLICANT: TAKAKURA, Hikaru  
APPLICANT: MORISHITA, Mio  
APPLICANT: SHIMOTO, Tomoko  
APPLICANT: ASADA, Kiyozo  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE  
FILE REFERENCE: TAKAKURA-6  
CURRENT APPLICATION NUMBER: US/09/445,472  
CURRENT FILING DATE: 1999-12-06  
PRIOR APPLICATION NUMBER: 151969/1997  
PRIOR FILING DATE: 1997-06-10  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: Patent In version 3.0

SEQ ID NO 12  
LENGTH: 659  
TYPE: PRT  
ORGANISM: Thermococcus celer  
US-09-445-472-12

Query Match 7.2%; Score 81; DB 4; Length 659;  
Best Local Similarity 25.9%; Pred. No. 4.2;  
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;

QY 41 KYODAK-----SIOGVSV--TYVAKTYENDTSGW--GSOFKCLOVEIERKEDYT 91  
DB 436 KIDVYAKLFTGTVADKSGATHTFDVSGATFTATLYWDTGSSDIDLTYLPNGNEVDY 495  
QY 92 VTSVETFRNASSPIKRYNTE---TVKAVFOYGYKNIRNAIEYOV-----GGC- 136  
DB 496 YTAHYGPEK---VGYNPTAGTWYKVV---SYKGAN---YQVDVDSGSLSGSGGN 545  
QY 137 -----LNTDPLFTDGLCDVFPYFNADGCELMVKKSHYKHVPDCTFEFNVFC 187  
DB 546 PNEPNPNPTPTDTOTFT-GSYND--YMDTSD-----FTFMVNS 583  
QY 188 AKDKTYDI-----FNEECYNGEP 207  
DB 584 GATKTTGDLTFTSTYNDLDTLYDP 608

## RESULT 7

US-08-487-890A-5  
Sequence 5, Application US/08487890A  
Patent No. 5708149

## GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-5

Query Match 7.1%; Score 79.5; DB 1; Length 913;  
Best Local Similarity 21.9%; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAAATQAEITTSARAGENPLMAHEELCKYOD-AKMS-----IDQGVSVTVLAKTITY 64  
DB 593 YVDLGLGMRVDVSRNANESTI---SVGKFNFSNITGIVIKPTFEMLDSTRLS-TGR 646  
QY 65 EN-----DTSGMSGOSFKCLOVEIERKEE-----DYTVSVTFRNASS 103  
DB 647 RNSPFAEMYGRGKADTIVYIKFK---PETSROEFELAKGDFGNIEHFSNA--700  
QY 104 PIKRYNTETVKAVFQ-----YGYKNIRNAIEYOVGGGLNITDPLFTDGLCDVFX 155  
DB 701 ---YRNLIAPAEELSKNGTGGKNGYGNHNAQKLV---GVNITQOLDENG-----745  
QY 156 VPNADGCELMVKKSHYKHVPD--YCTFEFNVFCADRK 192  
DB 746 -----LW-----KRIPYGYATFAFNRVYKVDOK 769

## RESULT 8

US-08-478-435-5  
Sequence 5, Application US/08478435  
Patent No. 5922323

## GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,435  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-462 MIS:vg

TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 913 amino acids
:   TYPE: amino acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:
US-08-478-435-5

Query Match
Best Local Similarity 21.9%; Pred. No. 9.9;
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

Oy 11 FVSAALATQAEITTSKAGNPPLMAHEELGKYQD-AMKS-----IDGCVSTVYLAATTY 64
Db 593 YVDGLGMRVDVSRKANEIYI-----SVGKFNFSNWTGIVKPTLWLDLSTYLS-TGF 646
Oy 65 EN-----DTGSMGSOFKCLOVQEIERRKE-----DYTVTSVTFERNASS 103
Db 647 RNPSFAEYGRYGGKDTDYIGKFK-----PETSBNDFGLALGDFGNITSHFSNA-- 700
Oy 104 PIKYNYNTEYTKAVFO-----YGYKNIRNAIEYOGGGLNTDTLIFTDGLCDVFX 155
Db 701 ---YRNLIAPAEELSKNGTTGKNGYGHNAQNAKLV---GVNITQQLDFNG----- 745
Oy 156 VPNAOQCELMVKKSHKHPD--YCTFEVNFCAKDRK 192
Db 746 -----LW-----KRIPGYATFAYNRVKKDOK 769

RESULT 9
US-08-337-483-5
: Sequence 5, Application US/08337483
: Patent No. 5922562
: GENERAL INFORMATION:
:   APPLICANT: Loosmore, Sheena
:   APPLICANT: Harkness, Robin
:   APPLICANT: Schryvers, Anthony
:   APPLICANT: Chong, Pele
:   APPLICANT: Gray-Owen, Scott
:   APPLICANT: Yang, Yan-Ping
:   APPLICANT: Murdin, Andrew
:   APPLICANT: Klein, Michel
:   TITLE OF INVENTION: Transferrin Receptor Genes
:   NUMBER OF SEQUENCES: 147
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Slim & McBurney
:     STREET: Suite 701, 330 University Avenue
:     CITY: Toronto
:     STATE: Ontario
:     COUNTRY: Canada
:     ZIP: M5G 1R7
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: Patentln Release #1.0, Version #1.25
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/08/337,483
:     FILING DATE: 08-NOV-1994
:     CLASSIFICATION: 435
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Stewart, Michael I
:     REGISTRATION NUMBER: 24,973
:     REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb
:   TELECOMMUNICATION INFORMATION:
:     TELEPHONE: (416) 595-1155
:     TELEFAX: (416) 595-1163
:   INFORMATION FOR SEQ ID NO: 5:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 913 amino acids
:       TYPE: amino acid
```

```

: STRANDEDNESS: single
: TOPOLOGY: linear
:
US-08-337-483-5

Query Match
Best Local Similarity 21.9%; Pred. No. 9.9;
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

Oy 11 FVSAALATQAEITTSKAGNPPLMAHEELGKYQD-AMKS-----IDGCVSTVYLAATTY 64
Db 593 YVDGLGMRVDVSRKANEIYI-----SVGKFNFSNWTGIVKPTLWLDLSTYLS-TGF 646
Oy 65 EN-----DTGSMGSOFKCLOVQEIERRKE-----DYTVTSVTFERNASS 103
Db 647 RNPSFAEYGRYGGKDTDYIGKFK-----PETSBNDFGLALGDFGNITSHFSNA-- 700
Oy 104 PIKYNYNTEYTKAVFO-----YGYKNIRNAIEYOGGGLNTDTLIFTDGLCDVFX 155
Db 701 ---YRNLIAPAEELSKNGTTGKNGYGHNAQNAKLV---GVNITQQLDFNG----- 745
Oy 156 VPNAOQCELMVKKSHKHPD--YCTFEVNFCAKDRK 192
Db 746 -----LW-----KRIPGYATFAYNRVKKDOK 769

RESULT 10
US-08-478-373-5
: Sequence 5, Application US/08478373
: Patent No. 5922841
: GENERAL INFORMATION:
:   APPLICANT: Loosmore, Sheena
:   APPLICANT: Harkness, Robin
:   APPLICANT: Schryvers, Anthony
:   APPLICANT: Chong, Pele
:   APPLICANT: Gray-Owen, Scott
:   APPLICANT: Yang, Yan-Ping
:   APPLICANT: Murdin, Andrew
:   APPLICANT: Klein, Michel
:   TITLE OF INVENTION: Transferrin Receptor Genes
:   NUMBER OF SEQUENCES: 147
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Slim & McBurney
:     STREET: Suite 701, 330 University Avenue
:     CITY: Toronto
:     STATE: Ontario
:     COUNTRY: Canada
:     ZIP: M5G 1R7
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: Patentln Release #1.0, Version #1.25
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/08/478,373
:     FILING DATE: 07-JUN-1995
:     CLASSIFICATION: 435
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: US 08/337,483
:     FILING DATE: 08-NOV-1994
:     CLASSIFICATION: 435
:     PRIOR APPLICATION NUMBER: US 08/175,116
:     FILING DATE: 29-DEC-1993
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: US 08/148,968
:     FILING DATE: 08-NOV-1993
:     ATTORNEY/AGENT INFORMATION:
:       NAME: Stewart, Michael I
:       REGISTRATION NUMBER: 24,973
:       REFERENCE/DOCKET NUMBER: 1038-463 MIS:vg
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: (416) 595-1155
:       TELEFAX: (416) 595-1163
```

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-478-373-5

Query Match 7.1%; Score 79.5; DB 2; Length 913;  
Best Local Similarity 21.9%; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEITTSKAGENPLMAHELLKQYD-AWKS-----IDQGVSVTVYLAHTY 64  
DB 593 YVDLGLGMRVDVSRKANEETI-----SVGKFKFNSWNTGIVIKPTFMDLSYRLS-TGF 646  
QY 65 EN-----DTGSMGSOFKLOVQEIETKE-----DYTVSVTFPRNASS 103  
DB 647 RNSFPAEMTGMRRGKDTDYIGKFK-----PSTRNDEFLALKGDFGNIETSHFSNA-- 700  
QY 104 PIRYVNTETVKAFFQ-----YGYKNIRNAIEYQVGGILNTDTLITDGLCDVfy 155  
DB 701 ---YRNLIAPAEELSKNGTGTGKNGYGHNNQNAKLV-----GVNITQDLDENG----- 745  
QY 156 VPNAOCCELMWKSKHKHVPD--YCTFVNVFCADRK 192  
DB 746 -----LW-----KRIPYGMVATFAVNRVKVDOK 769

## RESULT 11

US-08-474-671-5  
Sequence 5, Application US/08474671  
Patent No. 6008326

## GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,671  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-465 MIS:V9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163

## INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-474-671-5

Query Match 7.1%; Score 79.5; DB 3; Length 913;  
Best Local Similarity 21.9%; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEITTSKAGENPLMAHELLKQYD-AWKS-----IDQGVSVTVYLAHTY 64  
DB 593 YVDLGLGMRVDVSRKANEETI-----SVGKFKFNSWNTGIVIKPTFMDLSYRLS-TGF 646  
QY 65 EN-----DTGSMGSOFKLOVQEIETKE-----DYTVSVTFPRNASS 103  
DB 647 RNSFPAEMTGMRRGKDTDYIGKFK-----PSTRNDEFLALKGDFGNIETSHFSNA-- 700  
QY 104 PIRYVNTETVKAFFQ-----YGYKNIRNAIEYQVGGILNTDTLITDGLCDVfy 155  
DB 701 ---YRNLIAPAEELSKNGTGTGKNGYGHNNQNAKLV-----GVNITQDLDENG----- 745  
QY 156 VPNAOCCELMWKSKHKHVPD--YCTFVNVFCADRK 192  
DB 746 -----LW-----KRIPYGMVATFAVNRVKVDOK 769

## RESULT 12

US-08-483-577A-5  
Sequence 5, Application US/08483577A  
Patent No. 6015688

## GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 160  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,577A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-511  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-483-577A-5

Query Match 7.18; Score 79.5; DB 3; Length 913;  
Best Local Similarity 21.98; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

OY 11 FVSAALTAQAEETSAKAGENPLMAHEELGKYOD-AWKS-----IDGVSVTYVLAQTTY 64  
DB 593 YVDGLGMRDYDSTKANESTI-----SVGFKFNSNTGIVIPTEMLDSTYLS-TGF 646  
OY 65 EN-----DTGWSGQFKCLOYOEIRKEE-----DYTVTSVTFERNASS 103  
DB 647 RNPFAEMYGWRGCKDIDYIGKFK-----PETSBNQDFGLAKGDFGNIEISHFSNA-- 700  
OY 104 PIKYNTETEVKAVQ-----YQKNIRNALEYOGGGLNTDILIFDGLCDVYF 155  
DB 701 ---YRNLIAPAEELSKNGTGGKNGYGHMAQNAKLV---GVNITADLDFNG----- 745  
OY 156 VPNAOCCELWVKSHYKHPD--YCTFEVNFVFCADRK 192  
DB 746 -----LM-----KRIPYGMTATFAIRKRVKXDK 769

RESULT 13  
US-08-897-438-5  
Sequence 5, Application US/08897438  
Patent No. 6262016  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 160  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/897,438  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/483,577  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-897-438-5

Query Match 7.18; Score 79.5; DB 3; Length 913;  
Best Local Similarity 21.98; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

OY 11 FVSAALTAQAEETSAKAGENPLMAHEELGKYOD-AWKS-----IDGVSVTYVLAQTTY 64  
DB 593 YVDGLGMRDYDSTKANESTI-----SVGFKFNSNTGIVIPTEMLDSTYLS-TGF 646  
OY 65 EN-----DTGWSGQFKCLOYOEIRKEE-----DYTVTSVTFERNASS 103  
DB 647 RNPFAEMYGWRGCKDIDYIGKFK-----PETSBNQDFGLAKGDFGNIEISHFSNA-- 700  
OY 104 PIKYNTETEVKAVQ-----YQKNIRNALEYOGGGLNTDILIFDGLCDVYF 155  
DB 701 ---YRNLIAPAEELSKNGTGGKNGYGHMAQNAKLV---GVNITADLDFNG----- 745  
OY 156 VPNAOCCELWVKSHYKHPD--YCTFEVNFVFCADRK 192  
DB 746 -----LM-----KRIPYGMTATFAIRKRVKXDK 769

RESULT 14  
US-08-637-654-5  
Sequence 5, Application US/08637654  
Patent No. 6358727  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena M  
APPLICANT: Harkness, Robin E  
APPLICANT: Schryvers, Anthony B  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew D  
APPLICANT: Klein, Michel H  
TITLE OF INVENTION: HAEMOPHILUS TRANSFERRIN RECEPTOR GENES  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,654  
FILING DATE: 05-AUG-1996  
CLASSIFICATION: 435

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA94/00616
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-637-654-5

Query Match
Best Local Similarity 21.9%; Score 79.5; DB 4; Length 913;
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEITTSKAGENPLNHEELCKYOD-AWKS-----IDGVSVTYVLAKTTY 64
DB 593 YVDLGLGMRVDSRTKANESTI-----SVGKFNFSWNTGIIVIKPTMLDLSYRLS-TGF 646
QY 65 EN-----DTGSMGSOFKCLQVQEIETKEE-----DYTVSVTFERNASS 103
DB 647 RNSFPAEMTGMRYGKGDITDVIYIGFK-----PISRNOEFLALKGDFGNIIEISFNSA-- 700
QY 104 PIKYVNTETVKAVFQ-----YGYKNIRNAIEYQVGGGLNTDTLLFTDELCDVFX 155
DB 701 ---YRNLIAPAEELSKNGTGTGKNGYGHNAQNAKLV---GVNITQOLDENG----- 745
QY 156 VPNAOCCELMWKSKHKHVPD--YCFEYFNVCARDK 192
DB 746 -----LW-----KRIPGYMTAFAYNRVAVKVDOK 769

RESULT 15
US-08-649-518-5
; Sequence 5, Application US/08649518
; Patent No. 6361779
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray, Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,518
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,577
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-649-518-5

Query Match
Best Local Similarity 21.9%; Score 79.5; DB 4; Length 913;
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEITTSKAGENPLNHEELCKYOD-AWKS-----IDGVSVTYVLAKTTY 64
DB 593 YVDLGLGMRVDSRTKANESTI-----SVGKFNFSWNTGIIVIKPTMLDLSYRLS-TGF 646
QY 65 EN-----DTGSMGSOFKCLQVQEIETKEE-----DYTVSVTFERNASS 103
DB 647 RNSFPAEMTGMRYGKGDITDVIYIGFK-----PISRNOEFLALKGDFGNIIEISFNSA-- 700
QY 104 PIKYVNTETVKAVFQ-----YGYKNIRNAIEYQVGGGLNTDTLLFTDELCDVFX 155
DB 701 ---YRNLIAPAEELSKNGTGTGKNGYGHNAQNAKLV---GVNITQOLDENG----- 745
QY 156 VPNAOCCELMWKSKHKHVPD--YCFEYFNVCARDK 192
DB 746 -----LW-----KRIPGYMTAFAYNRVAVKVDOK 769

Search completed: August 4, 2003, 17:18:16
Job time : 31 secs
```





XX WPI: 1998-018506/02.  
 DR N-PSDB; AAV00230.  
 XX  
 XX New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 XX  
 XX Example 2; Fig 4; 44pp; English.

XX This protein comprises tick Dermacentor reticulatus (Dr) novel  
 CC vasoactive amine binding protein (VABP) D.RET6. Its amino acid  
 CC sequence was deduced from a cDNA clone (see AAV00230) obtained  
 CC from a Dr salivary gland cDNA library. 3 Novel VABPs, designated  
 CC FS-HBP1, FS-HBP2 and MS-HBP1 (see AAV37446-48), of the tick  
 CC Rhipicephalus appendiculatus have also been identified. The  
 CC VABPs can be expressed in host cells using e.g. a baculovirus  
 CC expression system. They can be used: (i) to assay histamine (or  
 CC other VA such as serotonin) in body fluids or cell culture  
 CC supernatants, e.g. to monitor the effect of allergens; (ii) for  
 CC binding VA, e.g. to remove histamine from blood, food, cell  
 CC cultures etc.; (iii) as an antihistamine or anti-inflammatory  
 CC agents, e.g. for treating insect, snake or scorpion bites or  
 CC dermatitis, or as a carrier for slow release of histamine-related  
 CC compounds; (iv) in vaccines to protect against metazoan parasites,  
 CC especially in animals; (v) as reagents for studying inflammation,  
 CC involvement of VA in ulcer formation or the immune response etc.  
 CC VABPs provide a more sensitive assay for histamine than  
 CC low-affinity antibodies currently used. They may also be more  
 CC effective and safer than conventional antihistamines.  
 CC  
 XX Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 19; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-104;  
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMOVLLTFVSAALTAQAEFTTSKAGENPLMAHEELKGYDAMKSIDGVSVTYVLA 60  
 DB 1 MKMOVLLTFVSAALTAQAEFTTSKAGENPLMAHEELKGYDAMKSIDGVSVTYVLA 60  
 QY KTTENTGSMGSGFKLOVEIERKEDYTVTSVFERRNASSPIKYNTEYKAVFOY 120  
 DB KTTENTGSMGSGFKLOVEIERKEDYTVTSVFERRNASSPIKYNTEYKAVFOY 120  
 QY 121 GYKNIRNAIEYQVGGGINTDTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 DB 121 GYKNIRNAIEYQVGGGINTDTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 QY 181 FVENVFCADKDKRTYDIFNECVYNGEPWL 209  
 DB 181 FVENVFCADKDKRTYDIFNECVYNGEPWL 209

RESULT 2  
 AAY18081  
 ID AAY18081 standard; Protein; 209 AA.

XX AAY18081;  
 AC  
 XX  
 DE 06-AUG-1999 (first entry)  
 DT  
 XX  
 XX Histamine binding protein D.RET6..  
 KW Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 XX Rhipicephalus appendiculatus.  
 OS

XX WO9927104-A1.  
 PN  
 XX  
 XX 03-JUN-1999.  
 PD  
 XX  
 XX 26-NOV-1998; 98WO-GB03530.  
 PF  
 XX  
 XX 26-JUN-1998; 98GB-0013917.  
 PR  
 XX  
 XX 26-NOV-1997; 97GB-0025046.  
 PR  
 XX  
 XX (OXFO-) OXFORD VACS LTD.  
 PA  
 XX  
 XX Nuttall PA, Paesen GC;  
 PI  
 XX  
 XX WPI: 1999-357841/30.  
 DR  
 XX  
 XX N-PSDB; AAX76967.

PT Histamine and serotonin binding compounds useful for the treatment  
 of allergies  
 PT

Claim 1; Fig 4; 84pp; English.

XX This sequence is an example of a histamine or serotonin binding  
 CC compound (A), of the invention. The compounds are useful for regulating  
 CC the action of histamine and serotonin (in e.g. inflammation and gastric  
 CC acid secretion), the detection, quantification and removal of histamine  
 CC or serotonin (in animals, plants, cell cultures, food materials, or  
 CC humans) and in the treatment of various diseases and allergies  
 CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
 CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
 CC allergies, abnormal blood pressure, migraine, psychological disorders,  
 CC respiratory disease, and coronary heart disease). Histamine may also be  
 CC used to regulate cellular growth and tissue repair. The molecules may  
 CC also be used as components of vaccines directed against blood-sucking  
 CC ectoparasites.  
 CC  
 XX Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 20; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-104;  
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMOVLLTFVSAALTAQAEFTTSKAGENPLMAHEELKGYDAMKSIDGVSVTYVLA 60  
 DB 1 MKMOVLLTFVSAALTAQAEFTTSKAGENPLMAHEELKGYDAMKSIDGVSVTYVLA 60  
 QY KTTENTGSMGSGFKLOVEIERKEDYTVTSVFERRNASSPIKYNTEYKAVFOY 120  
 DB KTTENTGSMGSGFKLOVEIERKEDYTVTSVFERRNASSPIKYNTEYKAVFOY 120  
 QY 121 GYKNIRNAIEYQVGGGINTDTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 DB 121 GYKNIRNAIEYQVGGGINTDTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 QY 181 FVENVFCADKDKRTYDIFNECVYNGEPWL 209  
 DB 181 FVENVFCADKDKRTYDIFNECVYNGEPWL 209

RESULT 3  
 AAY18086  
 ID AAY18086 standard; Protein; 203 AA.

XX AAY18086;  
 AC  
 XX  
 DE 06-AUG-1999 (first entry)  
 DT  
 XX  
 XX Histamine binding protein Ih/Bm-HBP2.  
 KW Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW

KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS Boophilus microplus.  
 OS Ixodes hexagonus.  
 OS Synthetic.  
 PN WO9927104-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 XX 26-NOV-1998; 98WO-GB03530.  
 PF 26-JUN-1998; 98GB-0013917.  
 PR 26-NOV-1997; 97GB-0025046.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 DR WPI: 1999-357841/30.  
 DR N-PSDB; AAX76971.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment  
 of allergies  
 PS Claim 14; Fig 8; 84pp; English.  
 XX  
 CC This sequence is an example of a histamine or serotonin binding  
 compound (A), of the invention. cDNA encoding this sequence was isolated  
 from a mixed Boophilus microplus/Ixodes hexagonus cDNA expression  
 library. The compounds are useful for regulating the action of histamine  
 and serotonin (in e.g. inflammation and gastric acid secretion), the  
 detection, quantification and removal of histamine or serotonin (in  
 animals, plants, cell cultures, food materials, or humans) and in the  
 treatment of various diseases and allergies (e.g. type I hypersensitivity  
 reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic  
 dermatitis, insect bites and food and drug allergies, abnormal blood  
 pressure, migraine, psychological disorders, respiratory disease, and  
 coronary heart disease). Histamine may also be used to regulate cellular  
 growth and tissue repair. The molecules may also be used as components of  
 vaccines directed against blood-sucking ectoparasites.  
 CC  
 XX  
 SQ Sequence 203 AA:  
 Query Match 32.4%; Score 364; DB 20; Length 203;  
 Best Local Similarity 38.1%; Pred. No. 4e-28;  
 Matches 75; Conservative 38; Mismatches 78; Indels 6; Gaps 5;  
 OY 6 VLLTFVSAALATQAEETTSKAGENRPLMAHEBLCKGYDAMKSIDGVSVTYVLAKTYE 65  
 DB 8 VYLAIVTAADQAPSSSTRNEPLEKTYW-HNOTLGRYDAMKSIDGVTYVLRSTYN 66  
 OY 66 NDTGSMGQFCKLOVQEIKEEDYTVSVTFERNASSPIKYNTVETVKAFOYGYKNI 125  
 DB 67 NDS-VWGNFTCLSTVYSKESSTFTVEYNTTKNSQ--QWVSSENVTAQOEGY-SV 122  
 OY 126 RNAIEYOVGGGINTDPLIFTDGELCDVFPVPNADOGCELWVKSHYKHPDYCTFVENV 185  
 DB 123 KNIQMTTENNTKFNQDVTVFETDQCDVLIYIPYKEDGYELWVRSEYLTNPCCOFIDL 182  
 OY 186 FCAKDRKTYDIFNEECV 202  
 DB 183 -VALGRITYNISTPNCV 198

RESULT 4  
 AAY18085  
 ID AAY18085 standard; Protein; 203 AA.  
 XX  
 AC AAY18085;  
 XX  
 DT 06-AUG-1999 (first entry)

XX  
 DE Histamine binding protein; serotonin binding compound; inflammation;  
 XX gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS Boophilus microplus.  
 OS Ixodes hexagonus.  
 OS Synthetic.  
 PN WO9927104-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 XX 26-NOV-1998; 98WO-GB03530.  
 PF 26-JUN-1998; 98GB-0013917.  
 PR 26-NOV-1997; 97GB-0025046.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 DR WPI: 1999-357841/30.  
 DR N-PSDB; AAX76970.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment  
 of allergies  
 PS Claim 13; Fig 7; 84pp; English.  
 XX  
 CC This sequence is an example of a histamine or serotonin binding  
 compound (A), of the invention. cDNA encoding this sequence was isolated  
 from a mixed Boophilus microplus/Ixodes hexagonus cDNA expression  
 library. The compounds are useful for regulating the action of histamine  
 and serotonin (in e.g. inflammation and gastric acid secretion), the  
 detection, quantification and removal of histamine or serotonin (in  
 animals, plants, cell cultures, food materials, or humans) and in the  
 treatment of various diseases and allergies (e.g. type I hypersensitivity  
 reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic  
 dermatitis, insect bites and food and drug allergies, abnormal blood  
 pressure, migraine, psychological disorders, respiratory disease, and  
 coronary heart disease). Histamine may also be used to regulate cellular  
 growth and tissue repair. The molecules may also be used as components of  
 vaccines directed against blood-sucking ectoparasites.  
 CC  
 XX  
 SQ Sequence 203 AA:  
 Query Match 31.4%; Score 353; DB 20; Length 203;  
 Best Local Similarity 36.3%; Pred. No. 5.1e-27;  
 Matches 74; Conservative 41; Mismatches 81; Indels 8; Gaps 6;  
 OY 1 MKMOVLLITVSAALATQAEETTSKAG--ENPLMAHEBLCKGYDAMKSIDGVSVTYV 58  
 DB 1 MKALLIAGVYLAIVTAADQAPSSSTRNEPLEKTYW-HSKELKNQDAMKSIDGVSVTYV 59  
 OY 59 LAKTYENDTSGMGSQFCKLOVQEIKEEDYTVSVTFERNASSPIKYNTVETVKAFOY 118  
 DB 60 FLRSTYNNDS-VWGNFTCLSTVYSKESSTFTVEYNTTKNSQ--QWVSSTENVTAQO 116  
 OY 119 QYGYKRNINAEYOVGGGINTDPLIFTDGELCDVFPVPNADOGCELWVKSHYKHPDY 178  
 DB 117 EEGY-DVKNITQMTTENNTKFNQDVTVFETDQCDVLIYIPYKEDGYELWVRSEYLTNP 175  
 OY 179 CTFVENVFCADKRTYDIFNEECV 202  
 DB 176 CQFIDFL-VALGRITYNISTPNCV 198

```

RESULT 5
AAW37447
ID AAW37447 standard; Protein: 190 AA.
XX
AC AAW37447;
XX
DT 08-JUN-1998 (first entry)
XX
DE Tick vasoactive amine binding protein 2 FS-HBP2.
XX
KW Female-specific vasoactive amine binding protein 1; FS-HCPI;
KW histamine; serotonin; assay; antihistamine; anti-inflammatory;
KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;
KW transgenic animal; tick.
XX
OS Rhinicephalus appendiculatus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= Sig_peptide
XX
PN W09744451-A2.
XX
PD 27-NOV-1997.
XX
PE 19-MAY-1997; 97WO-GB01372.
XX
PR 18-APR-1997; 97GB-0007844.
PR 18-MAY-1996; 96GB-0010484.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI: 1998-018506/02.
DR N-PSDB; AAV00228.
XX
PT New vasoactive amine binding proteins and related nucleic acid,
PT vectors - transformed cells and transgenic animals, used for
PT assaying or removing histamine and as antihistamine or
PT anti-inflammatory agents
XX
PS Example 2; Fig 2; 44pp; English.
XX
CC This protein comprises tick Rhinicephalus appendiculatus (Ra) novel
CC female-specific histamine binding protein 2 (FS-HBP2). Its amino
CC acid sequence was deduced from a cDNA clone (see AAV00228) obtained
CC from a salivary gland cDNA library. FS-HBP1 (see AAW37446) and
CC male-specific HSP1 (see AAW37448) and a related protein, D.RET6 (see
CC AAW37449) from Dermacentor reticulatus, were also identified. These
CC novel vasoactive amine binding proteins (VABPs) can be expressed
CC in host cells using e.g. a baculovirus expression system. They can
CC be used: (i) to assay histamine (or other VA such as serotonin) in
CC body fluids or cell culture supernatants, e.g. to monitor the
CC effect of allergens; (ii) for binding VA, e.g. to remove histamine
CC from blood, food, cell cultures etc.; (iii) as an antihistamine or
CC anti-inflammatory agents, e.g. for treating insect, snake or
CC scorpion bites or dermatitis, or as a carrier for slow release of
CC histamine-related compounds; (iv) in vaccines to protect against
CC metazoan parasites, especially in animals; (v) as reagents for
CC studying inflammation, involvement of VA in ulcer formation or the
CC immune response etc. VABPs provide a more sensitive assay for
CC histamine than low-affinity antibodies currently used. They may
CC also be more effective and safer than conventional antihistamines.
XX
SQ Sequence 190 AA;
XX
Query Match 25.8%; Score 290.5; DB 19; Length 190;
Best Local Similarity 36.0%; Pred. No. 8.1e-21;
Matches 72; Conservative 32; Mismatches 79; Indels 17; Gaps 9;
XX
7 LLLTFVSAALATGAEITTSARAGENPLMAHELLGKTODAKNSIDGCVSYVYLAKTYEN 66
II: : : II | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 3 LLLISLALVLA-----LSQVKGNDPMDWADBAANGAHODANKSLKADVENYVYWKATYKN 57
Qy 67 DTGSMGSGFKCLOQVLEERKEEDTVTSVTF-RNASSPIKYYNVETFAVQYCKNI 125
Db 58 DP-VWGNDFTCVGMADVNDKESIOAEFLFNADJTNQF--ATEKVAIVKMYG-NR 113
Qy 126 RNAIEYQVGGILNTDPLIFPDGELCDVFPVNNAD---QGCCLMVKKSHKHVPDYCTFV 182
Db 114 ENAFRYTEDEGQVFTDVIAYSD--DNCVITYPGIDGNBEGYELMT--TDIDNIPANCLNR 170
Qy 183 FNVFCARDKRYTDFNEECV 202
Db 171 FNEY-AVGRETRDVFTSACL 189

RESULT 6
AAV18079
ID AAV18079 standard; Protein: 190 AA.
XX
AC AAV18079;
XX
DT 06-AUG-1999 (first entry)
XX
DE Histamine binding protein FS-HBP2.
XX
KW Histamine binding protein; serotonin binding compound; inflammation;
KW gastric acid secretion; allergy; type I hypersensitivity reaction;
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW respiratory disease; coronary heart disease; cellular growth regulator;
KW tissue repair; blood-sucking ectoparasite; therapy.
XX
OS Rhinicephalus appendiculatus.
XX
PN W09927104-A1.
XX
PD 03-JUN-1999.
XX
PE 26-NOV-1998; 98WO-GB03530.
XX
PR 26-JUN-1998; 98GB-0013917.
PR 26-NOV-1997; 97GB-0025046.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI: 1999-357841/30.
DR N-PSDB; AAX76965.
XX
PT Histamine and serotonin binding compounds useful for the treatment
PT of allergies
XX
PS Claim 1; Fig 2; 84pp; English.
XX
CC This sequence is an example of a histamine or serotonin binding
CC compound (A), of the invention. The compounds are useful for regulating
CC the action of histamine and serotonin (in e.g. inflammation and gastric
CC acid secretion), the detection, quantification and removal of histamine
CC or serotonin (in animals, plants, cell cultures, food materials, or
CC humans) and in the treatment of various diseases and allergies
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic
CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug
CC allergies, abnormal blood pressure, migraine, psychological disorders,
CC respiratory disease, and coronary heart disease). Histamine may also be
CC used to regulate cellular growth and tissue repair. The molecules may
CC also be used as components of vaccines directed against blood-sucking
CC ectoparasites.
XX
SQ Sequence 190 AA;
XX
Query Match 25.8%; Score 290.5; DB 20; Length 190;
Best Local Similarity 36.0%; Pred. No. 8.1e-21;

```





QY	120	YGYNNINNALEYOVGGGLNT--DTLIFTDELCDVFPV---NADOG-CELMYKSHYK	173
Db	107	Y--TTENGIKYVENQGRTRQTFEEVFWFSYDKNCVDYIFVPERGSDGDEYELWSEDKID	163
QY	174	HVPDYCFEVENVFCADKDKT	193
Db	164	KIPCCCKFTMAVFAQDEKT	183
RESULT 11			
ID	AAV18080	standard; Protein: 200 AA.	
AC	AAV18080;		
XX			
XX	06-AUG-1999	(first entry)	
XX			
DE		Histamine binding protein MS-HBPL.	
KW		Histamine binding protein; serotonin binding compound; inflammation;	
KW		gastric acid secretion; allergy; type I hypersensitivity reaction;	
KW		asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;	
KW		drug allergy; abnormal blood pressure; psychological disorder; vaccine;	
KW		respiratory disease; coronary heart disease; cellular growth regulator;	
KW		tissue repair; blood-sucking ectoparasite; therapy.	
XX			
OS		Rhhipcephalus appendiculatus.	
XX			
PN		W09927104-A1.	
PD		03-JUN-1999.	
XX			
XX		26-NOV-1998; 98WO-GDB03530.	
XX			
XX		26-JUN-1998; 98GB-0013917.	
PR		26-NOV-1997; 97GB-0025046.	
XX			
XX		(OXFO-) OXFORD VACS LTD.	
XX			
PI		Nuttall PA, Paesen GC;	
XX			
DR		WPI: 1999-357841/30.	
XX		N-PSDB; MAX76966.	
PT			
PT		Histamine and serotonin binding compounds useful for the treatment	
PT		of allergies	
PS			
XX		Claim 1; Fig 3; 84p; English.	
XX			
CC	This sequence is an example of a histamine or serotonin binding		
CC	compound (A), of the invention. The compounds are useful for regulating		
CC	the action of histamine and serotonin (in e.g. inflammation and gastric		
CC	acid secretion), the detection, quantification and removal of histamine		
CC	or serotonin (in animals, plants, cell cultures, food materials, or		
CC	humans) and in the treatment of various diseases and allergies		
CC	(e.g. type I hypersensitivity reactions, urticaria, asthma, allergic		
CC	rhinitis (hay fever), atopic dermatitis, insect bites and food and drug		
CC	allergies, abnormal blood pressure, migraine, psychological disorders,		
CC	respiratory disease, and coronary heart disease). Histamine may also be		
CC	used to regulate cellular growth and tissue repair. The molecules may		
CC	also be used as components of vaccines directed against blood-sucking		
CC	ectoparasites.		
XX			
XX			
Sequence	200 AA;		
Query Match	23.6%;	Score 265;	DB 20; Length 200;
Best Local Similarity	34.5%;	Pred. NO. 3.1e-18;	
Matches	69;	Conservative 34;	Mismatches 71; Indels 26; Gaps 10;
QY	3	MOVVLLTFTVSAALATQAEETTSAKAGEPPLNAHBEELGKYDQDAKMSIDGVSYYVLAKT	62
Db	1	IKVILLV--LCAALCONADA-----NPTVANEKRLSYQDAWKSLODDQKKRYLQVA	51

QY	63	TYENDGSGNSGSOFCLOV--OQIEKEDDYVTSSTFRNMS-SPIKYYNTEYVAVPO	119
Db	52	TQTTD-GVAGGEETCTSVTAERIGRRKLNMT----LYKNHLDLDESHETITVWKA	106
QY	120	YGKTNLRNIAIEVQVGGGLNIT--DPLIFDGLCDVFPV----NADQG-CELMWKS	173
Db	107	Y---TTENGKIKETQGTQRTQTFEDVFPVSDIKNCDDVITFVKERSDSDGDI	163
QY	174	HVPDYCTFVFNVCADKRRKT	193
Db	164	KIPDCCFTWAFYFACQOQEXT	183
RESULT 12			
AAW37446			
AAW37446		standard; Protein; 190 AA.	
AAW37446:			
08-JUN-1998		(first entry)	
DE		Tick vasoactive amine binding protein, 1 FS-HBP1.	
XX			
XX			
KW		Female-specific vasoactive amine binding protein 1; FS-HCP1;	
KW		histamine; serotonin; assay; antihistamine; anti-inflammatory;	
KM		insect bite; snake bite; scorpion bite; dermatitis; vaccine;	
KM		transgenic animal; tick.	
XX			
OS		Rhipicephalus appendiculatus.	
XX			
FH		Key	Location/Qualifiers
FT		Peptide	1..18
FT			/label= Sig_peptide
XX			
PN		W09744451-A2.	
XX			
PD		27-NOV-1997.	
XX			
PF		19-MAY-1997;	97WO-GB01372.
XX			
PR		18-APR-1997;	97GB-0007844.
PR		18-MAY-1996;	96GB-0010484.
XX			
PA		(OXFO-) OXFORD VACS LTD.	
XX			
PI		Nuttall PA, Paesen GC;	
XX			
DR		WPI: 1998-018506/02.	
DR		N-PSDB; AAV00227.	
PT		New vasoactive amine binding proteins and related nucleic acid,	
PT		vectors - transformed cells and transgenic animals, used for	
PT		assaying or removing histamine and as antihistamine or	
PT		anti-inflammatory agents	
XX			
PS		Example 2; Fig 1; 44pp; English.	
XX			
CC		This protein comprises tick Rhipicephalus appendiculatus (Ra) novel	
CC		female-specific histamine binding protein 1 (FS-HBP1). Its amino	
CC		acid sequence was deduced from a cDNA clone (see AAV00227) obtained	
CC		from a salivary gland cDNA library. FS-HSP2 and male-specific	
CC		HSP1 (see AAW37447-48) and a related protein, D.REN6 (see AAW37449)	
CC		from dermactenter reticularis, were also identified. These novel	
CC		vasoactive amine binding proteins (VABPs) can be expressed in	
CC		host cells using e.g. a baculovirus expression system. They can	
CC		be used: (i) to assay histamine (or other VA such as serotonin) in	
CC		body fluids or cell culture supernatants, e.g. to monitor the	
CC		effect of allergens; (ii) for binding VA, e.g. to remove histamine	
CC		from blood, food, cell cultures etc.; (iii) as an antihistamine or	
CC		anti-inflammatory agents, e.g. for treating insect, snake or	
CC		scorpion bites or dermatitis, or as a carrier for slow release of	
CC		histamine-related compounds; (iv) in vaccines to protect against	

CC metazoan parasites, especially in animals; (v) as reagents for  
CC studying inflammation, involvement of VA in ulcer formation or the  
CC immune response etc. VABPs provide a more sensitive assay for  
CC histamine than low-affinity antibodies currently used. They may  
CC also be more effective and safer than conventional antihistamines.

XX Sequence 190 AA;

Query Match 22.4%; Score 252; DB 19; Length 190;  
Best Local Similarity 33.5%; Pred. No. 5,7e-17;  
Matches 68; Conservative 28; Mismatches 89; Indels 18; Gaps 8;

QY 3 MOVVLLTFVSALATQAEFTTSKAGENPLMAHELLGKYDANKSIDQVSTYVLA KT 62  
D 1 MKLLSLAFVLAISQVKA-----DKPWADAEANGEHODAMKHLQKLEENYDLIKA 52  
QY 63 TYENDTSGMSQSFCLQVQEIETKEEDYTVTSVTFERNASSPIKYVNTETVKA VFOYGY 122  
D 53 TYKNDP-VKGNDFTCVGTAAQNLNEDENKVEAMFPMNNADTV-YQHTFEKATPDKMYGY 110  
QY 123 KNINNALEYOVGGGLNTDPLIFPDGELCDVFYVPNAD---QGCCLWVKKSHYKHPDYC 179  
D 111 -NKENAITQTEDEQVLTDLVAFSD-DNCYVITVAGPDGSGAGYELMA--TDYTDVPASC 166  
QY 180 TFVFNVCAMDRTYDIFNEECV 202  
D 167 LEKENEYAA-GLPVRDVTSDCL 188

## RESULT 13

AA18078  
ID AA18078 standard; Protein: 190 AA.

XX AAY18078;

DT 06-AUG-1999 (first entry)

XX Histamine binding protein FS-HBP1.

XX Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.

XX Rhinicephalus appendiculatus.

XX MO9927104-A1.

XX 03-JUN-1999.

XX 26-NOV-1998; 98MO-GB03530.

XX 26-JUN-1998; 98GB-0013917.

XX 26-NOV-1997; 97GB-0025046.

XX (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC.

XX WPI; 1999-357841/30.

XX N-PSDB; AAX76964.

PT Histamine and serotonin binding compounds useful for the treatment  
XX of allergies

PS Claim 1; Fig 1; 84pp; English.

XX This sequence is an example of a histamine or serotonin binding  
CC compound (A), of the invention. The compounds are useful for regulating  
CC the action of histamine and serotonin (in e.g. inflammation and gastric  
CC acid secretion), the detection, quantification and removal of histamine

CC or serotonin (in animals, plants, cell cultures, food materials, or  
CC humans) and in the treatment of various diseases and allergies  
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
CC allergies, abnormal blood pressure, migraine, psychological disorders,  
CC respiratory disease, and coronary heart disease). Histamine may also be  
CC used to regulate cellular growth and tissue repair. The molecules may  
CC also be used as components of vaccines directed against blood-sucking  
CC ectoparasites.

XX Sequence 190 AA;

Query Match 22.4%; Score 252; DB 20; Length 190;  
Best Local Similarity 33.5%; Pred. No. 5,7e-17;  
Matches 68; Conservative 28; Mismatches 89; Indels 18; Gaps 8;

QY 3 MOVVLLTFVSALATQAEFTTSKAGENPLMAHELLGKYDANKSIDQVSTYVLA KT 62  
D 1 MKLLSLAFVLAISQVKA-----DKPWADAEANGEHODAMKHLQKLEENYDLIKA 52  
QY 63 TYENDTSGMSQSFCLQVQEIETKEEDYTVTSVTFERNASSPIKYVNTETVKA VFOYGY 122  
D 53 TYKNDP-VKGNDFTCVGTAAQNLNEDENKVEAMFPMNNADTV-YQHTFEKATPDKMYGY 110  
QY 123 KNINNALEYOVGGGLNTDPLIFPDGELCDVFYVPNAD---QGCCLWVKKSHYKHPDYC 179  
D 111 -NKENAITQTEDEQVLTDLVAFSD-DNCYVITVAGPDGSGAGYELMA--TDYTDVPASC 166  
QY 180 TFVFNVCAMDRTYDIFNEECV 202  
D 167 LEKENEYAA-GLPVRDVTSDCL 188

## RESULT 14

AAB74290  
ID AAB74290 standard; Protein: 182 AA.

XX AAB74290;

DT 20-JUN-2001 (first entry)

XX Histacalin protein MS-HBP1.

XX Histacalin; MS-HBP1; conjunctivitis.

XX Unidentified.

XX WO200115719-A2.

XX 08-MAR-2001.

XX 24-AUG-2000; 2000MO-GB03282.

XX 01-SEP-1999; 99GB-0020674.

XX (EVOL-) EVOLUTEC LTD.

XX Nuttall PA, Paesen GC.

XX WPI; 2001-257675/26.

XX Use of histacalin proteins for treating or preventing non-infective  
PT conjunctivitis, or for manufacturing a medicament for treating or  
PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
XX conjunctivitis

XX Claim 4; Page 5-6; 19pp; English.

XX The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein FS-HBP1. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.





THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 52.3467 Seconds  
(without alignments)  
897.203 Million cell updates/sec

Title: US-10-087-195-8  
Perfect score: 989  
Sequence: 1 NPTWANEAKIGSYQDAWKSL.....EKTVRNYTDSCKPAPAQN 182

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacterioph:\*
- 17: sp\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	251.5	25.4	209	5	Q8WSK7 dermatocentor
2	113.5	11.5	330	3	Q05583 saccharomyc
3	108.5	11.0	311	5	Q8MVC2 ixodes scap
4	95.5	9.7	210	5	Q8MVC2 ixodes scap
5	92.5	9.3	220	5	Q8MVC2 ixodes scap
6	92.5	9.3	306	5	Q8MVC2 ixodes scap
7	92.5	9.3	581	4	Q8MVC2 ixodes scap
8	92.5	9.3	956	5	Q8MVC2 ixodes scap
9	87.5	8.8	169	2	Q05273 bacillus su
10	87.5	8.8	483	16	Q928X7 listeria mo
11	87.5	8.8	483	16	Q928X7 listeria mo
12	86.5	8.7	604	11	Q8K361 mus musculu
13	86.5	8.7	3467	5	Q81218 plasmodium
14	86.5	8.7	391	16	Q8FQ68 corynebacte
15	85.5	8.6	428	16	Q8F3F8 leptospira
16	85.5	8.6	1543	5	Q8MVC7 drosophila

17	85.5	8.6	1970	5	Q9VCH4	Ogvech drosophila
18	85.5	8.6	514	10	Q8L6A8	Q8L6A8 theobroma c
19	85.5	8.6	662	5	Q61096	Q61096 trypanosoma
20	85.5	8.6	1959	16	Q8YANI	Q8YANI listeria mo
21	85.5	8.6	582	10	Q43079	Q43079 plasm saliv
22	85.5	8.4	1073	5	Q81JWS	Q81JWS plasmodium
23	85.5	8.4	778	2	Q9EXN1	Q9EXN1 enterococcu
24	83.5	8.4	1078	12	Q90193	Q90193 agaricus bl
25	83.5	8.4	1124	5	Q8L1S3	Q8L1S3 plasmodium
26	83.5	8.4	3896	12	P87514	P87514 border dise
27	82.5	8.3	306	9	Q64070	Q64070 bacterioph
28	82.5	8.3	306	16	Q31951	Q31951 bacillus su
29	82.5	8.3	498	16	Q8F3P9	Q8F3P9 leptospira
30	82.5	8.3	1802	5	Q96170	Q96170 plasmodium
31	82.5	8.3	392	16	Q8EAC0	Q8EAC0 streptococc
32	82.5	8.3	392	16	Q8D1V2	Q8D1V2 streptococc
33	82.5	8.3	577	16	Q8E147	Q8E147 shewanella
34	82.5	8.3	617	16	Q8X116	Q8X116 clostridium
35	81.5	8.2	196	16	Q8R7C5	Q8R7C5 thermomater
36	81.5	8.2	388	16	Q8R1M4	Q8R1M4 fusobacteri
37	81.5	8.2	501	16	Q8PQC5	Q8PQC5 xanthomonas
38	81.5	8.2	817	16	Q93214	Q93214 staphylococ
39	81.5	8.2	223	5	Q9GP33	Q9GP33 echinococcu
40	81.5	8.2	581	6	Q95LY3	Q95LY3 macaca fasc
41	81.5	8.2	912	2	Q06915	Q06915 lactobacill
42	81.5	8.2	953	16	Q8EVP8	Q8EVP8 mycoplasma
43	80.5	8.1	433	17	Q97AY2	Q97AY2 thermoplasma
44	80.5	8.1	641	5	Q05508	Q05508 trypanosoma
45	80.5	8.1	721	16	Q8Y763	Q8Y763 listeria mo

#### ALIGNMENTS

#### RESULT 1

ID	Q8WSK7	PRELIMINARY:	PRT:	209 AA.
AC	Q8WSK7:			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Serotonin and histamine binding protein.			
OS	Dermacentor reticulatus.			
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
OC	Parasitiformes; Ixodidae; Ixodidae; Dermacentor.			
OX	NCBI_TaxID=57047;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Saunders, G.C., Nuttall P.A.;			
RT	"A high affinity serotonin- and histamine-binding lipocalin secreted by blood-feeding ticks."			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF217101; AAL56644.1;			
DR	InterPro: IPR002970; His_binding.			
DR	Pfam: PF02098; His_binding; 1.			
SO	SEQUENCE 209 AA; 23889 MW; E5EB8BE710859591 CRC64;			

QY	Query Match	25.4%	Score 251.5;	DB 5;	Length 209;
DB	Best Local Similarity	35.3%	Pred. No. 2.1e-14;		
DB	Matches	61;	Conservative	29;	Mismatches 66;
DB					Indels 17;
DB					Gaps 8;
QY	1 NPTWANEAKIGSYQDAWKSLQDDONKRYVLAQQTTP-GWAGCEFTCVSYAEIKKK 59				
DB	30 NPLWAEHLGKRYDAMKSLDQGVTVVLAQTTTENDTGSWSQFKLV--QETIKKE 87				
QY	60 LNATL---LYKNKHLTDLKSHERITWKAADV---TTENGIRKETGTPTQPEDV 112				
DB	88 EDYITVSFTFRNMS-SPIKYVNTVEYKAVQYQYKIRNAIEQVGGGLNIT--DPLI 144				
QY	113 FSDYKNCVIFVPERGSDGDEYELWSEDKIDKIPDCCKFTMAVFAQQQEXT 165				
DB	145 FTDGLCDVFPVY---NADQG-CELMVKKSHKHPDYCTFVFNFVCAKDRKT 193				

RESULT 2  
ID 005583 PRELIMINARY; PRT; 330 AA.  
AC 005583;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Similar to S. CEREVISIAE SEC13P.  
GN YDR267C OR D9954.11.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Le T.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,  
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,  
RA Johnston L., Langston T., Lacroix P., Le T., Mardis E., Menezes S.,  
RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,  
RA Talch A., Trevaaskis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,  
RA Wilson R., Waterston R.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Waterston R.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Jia Y., Cherry J.M.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U51030; AAB64456.1; -  
DR SGD: S0002675; YDR267C.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 7.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR PRODOM: PD000018; WD40; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 5.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 330 AA; 37274 MW; B44CCAA3125FD666 CRC64;  
Query Match 11.5%; Score 113.5; DB 3; Length 330;  
Best Local Similarity 27.5%; Pred. No. 0.051;  
Matches 38; Conservative 17; Mismatches 62; Indels 21; Gaps 5;  
QY 16 AMKSLQDQDKRRYLAQATQTTDGVW-----GEFTCVSVTAERIGKRLNATILYKN 68  
DB 114 AM-----SNDGYLLATCSR-DKSVMIWETDESGEEYECISVQEH--SQDVKHVIMHPS 164  
QY 69 KHLTDKSHETITWKAYITTE-----NGIYETOGTITQTFEDVYFVSIDYKNCVIF 123  
DB 165 EALLASSSYDDTFVRIHMDYDDDWECVAVLNGHGTWSSDPTTEGVFRICSGSDSTVR 224  
QY 124 VPKRSGDEGDYELWASE 141  
DB 225 YMKYMGDDDDDDQEWVCE 242

AC 08MVC2;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
RA Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis."  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF483712; AAM93664.1; -  
SQ SEQUENCE 210 AA; 24689 MW; C5B75614649A9BED CRC64;  
Query Match 9.7%; Score 95.5; DB 5; Length 210;  
Best Local Similarity 22.4%; Pred. No. 1.1; Mismatches 59; Indels 43; Gaps 9;  
Matches 38; Conservative 30; Mismatches 59; Indels 43; Gaps 9;  
QY 7 EARLGSYDAMKSLQD-QDNKRYLAQATQTTDGVWGEFTCVS-----V 50  
DB 29 DEKTYMOYODIORLNNPDRESMLYIRYRETQ---GSEHICSAKASENQPGSDYEFV 85  
QY 51 TAERIGKKILN---ATILYKNKHLTDL---KESHETITWKAYDYTTENGIKYETQGT 103

Db 86 GERYLGTEKQNRKRTVTLVATPYKTEMHATQRONNANMVSQKDD--ADGKGXG----- 138  
QY 104 TTFEDVVFVSDYKNCDFVFKRSGDEGDEYELWVSEDKIDK-IPDCK 152  
139 -----LIYSDYNKCDILRLVRENSGH--DCELYLHRSKALDDGVPRECE 179

RESULT 5  
ID Q8MVB8 PRELIMINARY: PRT; 220 AA.  
AC Q8MVB8; 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DE Putative secreted protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rhode Island; TISSUE=Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
RT Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis";  
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF483722; AAM93644.1; -  
SQ SEQUENCE 220 AA; 25375 MW; 61ADD7258AA28DDE CRC64;

Query Match 9.3%; Score 92; DB 5; Length 220;  
Best Local Similarity 23.7%; Pred. No. 2.5;  
Matches 47; Conservative 29; Mismatches 70; Indels 52; Gaps 12;

QY 6 NEAKLSYODAMKSLQODQNKRYLAQATQTDGVWGEFFCVSYTA----- 52  
Db 46 NSAKL-----WL-----KRRKFLYVRSKANDTEGCGACQVOKKAFIELPKRFSWN 94  
QY 53 EIKGKKLNLATILYKKNHLDLTKESHETTVKAVDYTTENGICKETG---GTPTQT--- 106  
Db 95 TPVTLKLVFSSVSGYRNQTHNDGFTALITTKSYE-----GLVYDDMFSGIELTYTHYI 149  
QY 107 FEDV-FVFSDDYKNCDFVFKRSGDEGDEYELWVSEDK--IDKIPDCKFTMAYFAQDQ 163  
Db 150 YVDYELFLFDYETCTTMRP---SDD-VYVMWMDIRNITDINPQC---ETAYHGPNVE 201  
QY 164 -----KTVRNVTDSQC 175  
Db 202 YGCTVWPKRYVYDDMIC 219

RESULT 6  
ID Q8MVC3 PRELIMINARY: PRT; 306 AA.  
AC Q8MVC3; 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DE Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rhode Island; TISSUE=Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
RT Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis";  
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF483717; AAM93639.1; -

SQ SEQUENCE 306 AA; 34232 MW; B2DB09181CC4395D CRC64;

Query Match 9.3%; Score 92; DB 5; Length 306;  
Best Local Similarity 27.3%; Pred. No. 3.7;  
Matches 42; Conservative 25; Mismatches 61; Indels 26; Gaps 10;

QY 1 NPTWANEAKLSYODAMKSLQODQNKRYLAQATQTDGVWGEFFCVSYTAEKICK--K 58  
Db 30 NPL-LNPSLGLPQSAKAKINSKSDPFVLMFRSRNH-----PNTCVVVTASLNETLK 84  
QY 59 KLNATILYKKNHLDLTKES--HETTVKAVDYTTENGICKYETGTF---RTQTEDEVFV 112  
Db 85 IYNFRTTYNE--TTGKNNTLEYQYRALNQTYKLEENVRAGLKGTPSDKPPPLSGNMVI 142  
QY 113 -FSDYKNC-----DVIFPKER-GSDEGDEY 136  
Db 143 EYGDV-SCNTSSKRPYDMLAKKADAVGSEAGSAE 175

RESULT 7  
ID Q8NA03 PRELIMINARY: PRT; 581 AA.  
AC Q8NA03; 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DE Hypothetical protein FLJ35989.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,  
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,  
RA Matsumura Y., Moriya S., Chiba E., Nomura H., Onogawa S.,  
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,  
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,  
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,  
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hiro Y., Saito K., Nishikawa T.,  
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,  
RA Kikuchi H., Kanda K., Magatsuma M., Murakawa K., Kanehori K.,  
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,  
RA Sugano S., Nagahara K., Masuno Y., Nagai A., Isogai T.;  
RT "NEO human cDNA sequencing project";  
RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AK093308; BAC04128.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 581 AA; 66121 MW; 979A8B4DE2177C74 CRC64;

Query Match 9.3%; Score 92; DB 4; Length 581;  
Best Local Similarity 25.2%; Pred. No. 8;  
Matches 40; Conservative 25; Mismatches 66; Indels 28; Gaps 7;

QY 4 WANEAKLSYODAMKSLQODQNKRYLAQATQTDGVWGEFFCVSYTAEKICKRL 60  
Db 151 W-EIKSAKYSSEAWOSKEKMENTKFLSLTAAVESEVGPSSHEEDTFSSVFTQIIPPEY 209  
QY 61 NATIILYKKNHLDLTKESHETTVKAVDYTTENGICKYETGTF---EDVVFSDYK 117  
Db 210 ENQOMKLNDFPCVDVRNLSLKSGKPPSNTF---KILRKHNDQFIKRIELAKESR 266  
QY 118 NCDVIFPKER-----GSDEGDEYLV 139  
Db 267 N-PVWVDREKRLVELLKLDDEKDSGLSSSEGDSGW 304

RESULT 8  
ID Q8IB10 PRELIMINARY: PRT; 956 AA.  
AC Q8IB10; 01-MAR-2003 (TREMBLrel. 23, Created)

01-MAR-2003 (TrEMBLrel. 23, last sequence update)  
01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
Hypothetical protein.  
GN PF07\_0114.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,  
RA Quail M., Barrall B.,  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: A844506; CAD51020.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 956 AA; 114829 MW; BF81CA983BF55B0A CRC64;  
Query Match 9.3%; Score 92; DB 5; Length 956;  
Best Local Similarity 28.2%; Pred. No. 15;  
Matches 33; Conservative 13; Mismatches 35; Indels 36; Gaps 6;  
QY 58 KKLNAATLYKNKHLTDI--KESHTTVMKAYDYTTENGIKYGTGTGTGTFEDVFSVSD 115  
DB 24 KKYNEFFYKKNKLTLDKSKRENNHFDIYKLY-----EKKNYITD 63  
QY 116 YKNCQVTFVPERGSDGDEYELWVSEDKIDKIPDCKFTMAYFAQOQEKTVRNYTID 172  
DB 64 Y--CNLTV-----DEIDNSFKIEDN-KSHINICK---YFLQNFDEVTNNYTD 106  
RESULT 9  
005273 PRELIMINARY; PRT; 169 AA.  
AC 005273;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)  
DE Hypothetical 19.8 kDa protein.  
GN YUCC.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA SPRATN-168;  
RA Oudega B., Koningsstein G., Duisterhoef A.;  
RT Bacillus subtilis genome project, DNA sequence from yucca to yuch.;"  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z33940; CAB07966.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 169 AA; 19833 MW; 211319D174AF435F CRC64;  
Query Match 8.8%; Score 87; DB 2; Length 169;  
Best Local Similarity 26.9%; Pred. No. 4.9;  
Matches 35; Conservative 19; Mismatches 48; Indels 28; Gaps 6;  
QY 45 FT-CVSTAEKIKGKRLNA--TILYKNKHLTDK-----ESHETITVMKAYDYTTENG 94  
DB 26 FTWCQHLTTEELKLNKLGVEVILLYFLFIIDVEYSWIRALQCKEDIAVOFA-DYOTLTK 84  
QY 95 IKYETGTGTQTFEDVFSVSEDKNCVDFVPERGSD-----SDYEL 137  
DB 85 VASLSMTFTREITIDVQTHSDQIKDELVSVPWETGVLYTRDELHHTIHAHEIHGOLSV 144  
QY 138 WVEDKIDKI 147  
DB 145 WARELKLSPV 154  
RESULT 10  
092BX7 PRELIMINARY; PRT; 483 AA.  
AC 092BX7;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

01-DEC-2001 (TrEMBLrel. 19, last sequence update)  
01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
Two-component sensor histidine kinase.  
GN LISK OR L1N1415.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-CLIP 11262 / Serovar 6a;  
RX PubMed=11679669;  
RA Glaser P., Franzeul L., Buchrieser C., Rusniok C., Amend A.,  
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouni F., Couve E., de Darivar A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Entian K.-D., Fahl H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kutznat G.,  
RA Madueno E., Maltounan A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;  
RT "Comparative genomics of Listeria species.";  
RL Science 294:849-852 (2001).  
DR EMBL: AL596168; CAC96646.1; -  
DR L1stList; L1N01415; -  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003660; HAM.  
DR InterPro: IPR003661; His\_kinA.  
DR InterPro: IPR005467; His\_kinase.  
DR Pfam: PF00672; HAM; 1.  
DR Pfam: PF02518; HAMase\_C; 1.  
DR Pfam: PF00512; HSKA; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR SMART: SM00304; HAM; 1.  
DR SMART: SM00387; HAMase\_C; 1.  
DR SMART: SM00388; HSKA; 1.  
DR PROSITE: PS50109; HIS\_KIN; 1.  
KM Kinase; Complete proteome.  
SQ SEQUENCE 483 AA; 55501 MW; F5712DF5304C32D2 CRC64;  
Query Match 8.8%; Score 87; DB 16; Length 483;  
Best Local Similarity 24.0%; Pred. No. 18;  
Matches 37; Conservative 29; Mismatches 58; Indels 30; Gaps 7;  
QY 18 KSLQODONKRYVLAQATQTTDGVWGEFTCVSTAEIKGKKLNATLYKNKHLTDKES 77  
DB 308 KKLVDENLDSRAEQISQTE---LQITVNAIVEQY---RRNFVMTENFTFT-LKED 359  
QY 78 HETITVMKAYDY-----TTENGIKYGTGTGTGTFEDVFSVSEDKNCVDFVPERGS 130  
DB 360 DTDLRALDQNNHLEQLIITIDNNAVYSGDGEV---DMHYVKEQKQNIH----- 406  
QY 131 DEGDYELWVSEDKIDKIPDCKFTMAYFAQOQEK 164  
DB 407 DVRDYGEGISQIEIDKIFN--RRYVDKARSREK 438  
RESULT 11  
09RPY9 PRELIMINARY; PRT; 483 AA.  
AC 09RPY9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
DE LISK (Two-component sensor histidine kinase).  
GN LISK OR LMO1378.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1639;  
RN [1]  
RP SEQUENCE FROM N.A.



DB 1375 EIMEGMLCALTHEIDEEBKKIKSTYSYDOLAKTTNGTPTLEKFAERQFIARWETENSDE 1434  
OY 45 FTCSVTAETKIGKKLNATILYKNNHLDLTKESHETITWMAV-DYTTENGICYETOCTR 103  
DB 1435 FCREREREKEVEYERCKKDHGCKNKPNT-KGNHCVSACADYEYISTKCKQYNTQ--- 1489  
OY 104 TQTFEDVFVPSDYKNCVDYFVPERKSGSEGDYELAWSDDKIDKIPD-CCKTMAVFAOQ 162  
DB 1490 KKKD-----IDKKNKNGEE-YENYKDKAEADYLDKDCFPQTCDY---M 1528  
OY 163 EKVTRN-----VYTD-----SCKPAP 179  
DB 1529 EKVKNSEYMDKPNKTYTNSDLEKKCKECPQ 1560  
RESULT 14  
O8F068  
ID O8F068 PRELIMINARY; PRT; 391 AA.  
AC O8F068:  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN CE1265.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Y5-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RA Kewarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,  
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,  
RA Usuda Y., Sugimoto S.;  
RT "The entire genomic sequence of Corynebacterium efficiens Y5-314.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP005218; BAC18075.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 391 AA; 42871 MW; 07C46C6DE4F41F0 CRC64;  
Query Match 8.7%; Score 86; DB 16; Length 391;  
Best Local Similarity 27.8%; Pred. No. 17;  
Matches 40; Conservative 15; Mismatches 43; Indels 46; Gaps 7;  
OY 27 RYVLAQTQTTD-----GYMGE--EFTCVSVTAEK-----IGKKLNATILYKN 68  
DB 218 RVWESGAPETTDHDLPTRVGGEVNSGNLVYVASKGHAVEQVPGKWFEEAL-----N 273  
OY 69 KHLTDKESHETITWMAVYTTENGICYETOCTR-----QTFEDVFVPSDYK 117  
DB 274 RELTSDIEDIEFTLTTRAYPDKSTYIKYGLQGTITLQNNRRLREGLASDEDIF----- 327  
OY 118 NCDVIFVPERKSGSD---EGDYEL 137  
DB 328 ---ASLKPRERTDVLVLEPQDEEL 348  
RESULT 15  
O8F3F8  
ID O8F3F8 PRELIMINARY; PRT; 428 AA.  
AC O8F3F8:  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Putative outer membrane protein.  
GN LA2448.  
OS Leptospira interrogans.  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
OX NCBI\_TaxID=173;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-56601 / Serogroup Icterohaemorrhagiae / Serovar 1a1;  
RA Ren S.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE011413; AA09647.1; -  
KW Complete proteome.  
SQ SEQUENCE 428 AA; 48954 MW; 0E7DD095D1BA1C7C CRC64;  
Query Match 8.6%; Score 85.5; DB 16; Length 428;  
Best Local Similarity 22.9%; Pred. No. 21;  
Matches 38; Conservative 31; Mismatches 60; Indels 37; Gaps 6;  
OY 5 ANEAKLSYODAMKSLQODNKRYYLAQATQTTGVMGEFTCVSVTAETKIGKKLNATI 64  
DB 29 ADEVKPGTYMDLTALQNPUNVRVL-----DLGGQNTTLPKEIEQL--KNLQKLY 77  
OY 65 LYKN-----KHLTDKESHETITWMAVYTTENGICYETOCTR----- 102  
DB 78 LFDNRRLKTLPEIKQLNQLNELSSNQLTLPKEIKLENLQRLDLDYDNRLLTLPPIEIG 137  
OY 103 RTQTFEDVFVPSDYKNCVDYFVPERKSGSEGDYELAWSDDKIDKIP 148  
DB 138 KLQMLQTLVLISSN---QTTLPRESGKLENLQELNLSNQLTLP 179

Search completed: August 1, 2003, 12:20:01  
Job time: 54.3467 secs



Accession	Length	Score	Expect	Ident	Accession	Length	Score	Expect	Ident
33	76	7.7	397	1	YBP2_YEAST	121	10.0	1.0	100
					P38226	121	10.0	1.0	100
					saccharomy	121	10.0	1.0	100

13 INFLUENCE OF VIBRATION ON THE STABILITY OF THE / 8

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525





Query Match 8.7%; Score 86; DB 1; Length 529;  
Best Local Similarity 21.4%; Pred. No. 4.2;  
Matches 42; Conservative 26; Mismatches 64; Indels 64; Gaps 8;

QY 5 ANEAKLSYODAWMSLOODNKKRYLLAQTOTGTGVEEFTCVSAERKGGKKNLTI 64  
DB 266 ATRAKTPPIYQIMRAFKNLED-----CESFTSG----- 294  
QY 65 LYNNKHLTDKESHETITWKAYDYTTENGIKYETOGTRQTFEDVEFSDKNCVIFV 124  
DB 295 -YSRPSPTLKADEFTKAIK--NYVQKREKVD--QATDKE-----DDYKMLDKIF- 341  
QY 125 PKRGSDEGYELWVSDKIDKIPDCKF-----TMATFAQOQKTVR 167  
DB 342 --KDKGFYPOKLDADKKDLKDATQTNKIKLADITDRSELKVLVYTRQKEOTLT 399  
QY 168 NYTDSCKPAPA-QN 182  
DB 400 KELEKAEKATQANQN 415

RESULT 6  
ID 2180 HUMAN STANDARD; PRT; 692 AA.  
AC 09JUG: 09PIU2.  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Zinc finger protein 180 (HH2168).  
GN ZNF180.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shannon M., Branscomb E., Hauser L., Gordon L., Ashworth L.,  
RA Strubs L.,  
RT "Differential expansion of homologous zinc-finger gene families in  
RT human chromosome 19q13.2 and mouse chromosome 7.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kdojiamani V., Ge Y., Krummel G., Grable L., Severin J., Gordon L.,  
RA Shannon M., Brower A., Olsen A.S., Smith L.M.,  
RT Sequence analysis of a 1Mb region in 19q13.2 containing a zinc finger  
RT gene cluster.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
CC FINGER PROTEINS.  
CC -1- SIMILARITY: Contains 1 KRAB domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
DR EMBL; AF192813; AAF07950.1;  
DR EMBL; AC069278; AAF71790.1;  
DR HSSP; P08047; ISP2.  
DR Genew; HGNC:12970; ZNF180.  
DR MIM: 606740;  
DR InterPro; IPR00109; KRAB.  
DR InterPro; IPR007087; ZNF\_C2H2.  
DR Pfam; PF01352; KRAB; 1.  
DR Pfam; PF00096; zf-C2H2; 12.  
DR Prodom; PD000003; Znf.C2H2; 11.  
DR SMART; SM00349; KRAB; 1.

DR SMART; SM00355; Znf.C2H2; 12.  
DR PROSITE; PS50805; KRAB; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 12.  
DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 12.  
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
KW Nuclear protein; Repeat.  
FT DOMAIN 72 145  
FT ZN\_FING 353 375 KRAB.  
FT ZN\_FING 381 403 C2H2-TYPE.  
FT ZN\_FING 409 431 C2H2-TYPE.  
FT ZN\_FING 437 459 C2H2-TYPE.  
FT ZN\_FING 465 487 C2H2-TYPE.  
FT ZN\_FING 493 515 C2H2-TYPE.  
FT ZN\_FING 521 543 C2H2-TYPE.  
FT ZN\_FING 549 571 C2H2-TYPE.  
FT ZN\_FING 577 599 C2H2-TYPE.  
FT ZN\_FING 605 627 C2H2-TYPE.  
FT ZN\_FING 633 655 C2H2-TYPE.  
FT ZN\_FING 661 683 C2H2-TYPE.  
FT ZN\_FING 692 714 C2H2-TYPE.  
FT ZN\_FING 721 743 C2H2-TYPE.  
FT ZN\_FING 751 773 C2H2-TYPE.  
FT ZN\_FING 781 803 C2H2-TYPE.  
FT ZN\_FING 811 833 C2H2-TYPE.  
FT ZN\_FING 841 863 C2H2-TYPE.  
FT ZN\_FING 871 893 C2H2-TYPE.  
FT ZN\_FING 901 923 C2H2-TYPE.  
FT ZN\_FING 931 953 C2H2-TYPE.  
FT ZN\_FING 961 983 C2H2-TYPE.  
FT ZN\_FING 991 1013 C2H2-TYPE.  
FT ZN\_FING 1021 1043 C2H2-TYPE.  
FT ZN\_FING 1051 1073 C2H2-TYPE.  
FT ZN\_FING 1081 1103 C2H2-TYPE.  
FT ZN\_FING 1111 1133 C2H2-TYPE.  
FT ZN\_FING 1141 1163 C2H2-TYPE.  
FT ZN\_FING 1171 1193 C2H2-TYPE.  
FT ZN\_FING 1201 1223 C2H2-TYPE.  
FT ZN\_FING 1231 1253 C2H2-TYPE.  
FT ZN\_FING 1261 1283 C2H2-TYPE.  
FT ZN\_FING 1291 1313 C2H2-TYPE.  
FT ZN\_FING 1321 1343 C2H2-TYPE.  
FT ZN\_FING 1351 1373 C2H2-TYPE.  
FT ZN\_FING 1381 1403 C2H2-TYPE.  
FT ZN\_FING 1411 1433 C2H2-TYPE.  
FT ZN\_FING 1441 1463 C2H2-TYPE.  
FT ZN\_FING 1471 1493 C2H2-TYPE.  
FT ZN\_FING 1501 1523 C2H2-TYPE.  
FT ZN\_FING 1531 1553 C2H2-TYPE.  
FT ZN\_FING 1561 1583 C2H2-TYPE.  
FT ZN\_FING 1591 1613 C2H2-TYPE.  
FT ZN\_FING 1621 1643 C2H2-TYPE.  
FT ZN\_FING 1651 1673 C2H2-TYPE.  
FT ZN\_FING 1681 1703 C2H2-TYPE.  
FT ZN\_FING 1711 1733 C2H2-TYPE.  
FT ZN\_FING 1741 1763 C2H2-TYPE.  
FT ZN\_FING 1771 1793 C2H2-TYPE.  
FT ZN\_FING 1801 1823 C2H2-TYPE.  
FT ZN\_FING 1831 1853 C2H2-TYPE.  
FT ZN\_FING 1861 1883 C2H2-TYPE.  
FT ZN\_FING 1891 1913 C2H2-TYPE.  
FT ZN\_FING 1921 1943 C2H2-TYPE.  
FT ZN\_FING 1951 1973 C2H2-TYPE.  
FT ZN\_FING 1981 2003 C2H2-TYPE.  
FT ZN\_FING 2011 2033 C2H2-TYPE.  
FT ZN\_FING 2041 2063 C2H2-TYPE.  
FT ZN\_FING 2071 2093 C2H2-TYPE.  
FT ZN\_FING 2101 2123 C2H2-TYPE.  
FT ZN\_FING 2131 2153 C2H2-TYPE.  
FT ZN\_FING 2161 2183 C2H2-TYPE.  
FT ZN\_FING 2191 2213 C2H2-TYPE.  
FT ZN\_FING 2221 2243 C2H2-TYPE.  
FT ZN\_FING 2251 2273 C2H2-TYPE.  
FT ZN\_FING 2281 2303 C2H2-TYPE.  
FT ZN\_FING 2311 2333 C2H2-TYPE.  
FT ZN\_FING 2341 2363 C2H2-TYPE.  
FT ZN\_FING 2371 2393 C2H2-TYPE.  
FT ZN\_FING 2401 2423 C2H2-TYPE.  
FT ZN\_FING 2431 2453 C2H2-TYPE.  
FT ZN\_FING 2461 2483 C2H2-TYPE.  
FT ZN\_FING 2491 2513 C2H2-TYPE.  
FT ZN\_FING 2521 2543 C2H2-TYPE.  
FT ZN\_FING 2551 2573 C2H2-TYPE.  
FT ZN\_FING 2581 2603 C2H2-TYPE.  
FT ZN\_FING 2611 2633 C2H2-TYPE.  
FT ZN\_FING 2641 2663 C2H2-TYPE.  
FT ZN\_FING 2671 2693 C2H2-TYPE.  
FT ZN\_FING 2701 2723 C2H2-TYPE.  
FT ZN\_FING 2731 2753 C2H2-TYPE.  
FT ZN\_FING 2761 2783 C2H2-TYPE.  
FT ZN\_FING 2791 2813 C2H2-TYPE.  
FT ZN\_FING 2821 2843 C2H2-TYPE.  
FT ZN\_FING 2851 2873 C2H2-TYPE.  
FT ZN\_FING 2881 2903 C2H2-TYPE.  
FT ZN\_FING 2911 2933 C2H2-TYPE.  
FT ZN\_FING 2941 2963 C2H2-TYPE.  
FT ZN\_FING 2971 2993 C2H2-TYPE.  
FT ZN\_FING 3001 3023 C2H2-TYPE.  
FT ZN\_FING 3031 3053 C2H2-TYPE.  
FT ZN\_FING 3061 3083 C2H2-TYPE.  
FT ZN\_FING 3091 3113 C2H2-TYPE.  
FT ZN\_FING 3121 3143 C2H2-TYPE.  
FT ZN\_FING 3151 3173 C2H2-TYPE.  
FT ZN\_FING 3181 3203 C2H2-TYPE.  
FT ZN\_FING 3211 3233 C2H2-TYPE.  
FT ZN\_FING 3241 3263 C2H2-TYPE.  
FT ZN\_FING 3271 3293 C2H2-TYPE.  
FT ZN\_FING 3301 3323 C2H2-TYPE.  
FT ZN\_FING 3331 3353 C2H2-TYPE.  
FT ZN\_FING 3361 3383 C2H2-TYPE.  
FT ZN\_FING 3391 3413 C2H2-TYPE.  
FT ZN\_FING 3421 3443 C2H2-TYPE.  
FT ZN\_FING 3451 3473 C2H2-TYPE.  
FT ZN\_FING 3481 3503 C2H2-TYPE.  
FT ZN\_FING 3511 3533 C2H2-TYPE.  
FT ZN\_FING 3541 3563 C2H2-TYPE.  
FT ZN\_FING 3571 3593 C2H2-TYPE.  
FT ZN\_FING 3601 3623 C2H2-TYPE.  
FT ZN\_FING 3631 3653 C2H2-TYPE.  
FT ZN\_FING 3661 3683 C2H2-TYPE.  
FT ZN\_FING 3691 3713 C2H2-TYPE.  
FT ZN\_FING 3721 3743 C2H2-TYPE.  
FT ZN\_FING 3751 3773 C2H2-TYPE.  
FT ZN\_FING 3781 3803 C2H2-TYPE.  
FT ZN\_FING 3811 3833 C2H2-TYPE.  
FT ZN\_FING 3841 3863 C2H2-TYPE.  
FT ZN\_FING 3871 3893 C2H2-TYPE.  
FT ZN\_FING 3901 3923 C2H2-TYPE.  
FT ZN\_FING 3931 3953 C2H2-TYPE.  
FT ZN\_FING 3961 3983 C2H2-TYPE.  
FT ZN\_FING 3991 4013 C2H2-TYPE.  
FT ZN\_FING 4021 4043 C2H2-TYPE.  
FT ZN\_FING 4051 4073 C2H2-TYPE.  
FT ZN\_FING 4081 4103 C2H2-TYPE.  
FT ZN\_FING 4111 4133 C2H2-TYPE.  
FT ZN\_FING 4141 4163 C2H2-TYPE.  
FT ZN\_FING 4171 4193 C2H2-TYPE.  
FT ZN\_FING 4201 4223 C2H2-TYPE.  
FT ZN\_FING 4231 4253 C2H2-TYPE.  
FT ZN\_FING 4261 4283 C2H2-TYPE.  
FT ZN\_FING 4291 4313 C2H2-TYPE.  
FT ZN\_FING 4321 4343 C2H2-TYPE.  
FT ZN\_FING 4351 4373 C2H2-TYPE.  
FT ZN\_FING 4381 4403 C2H2-TYPE.  
FT ZN\_FING 4411 4433 C2H2-TYPE.  
FT ZN\_FING 4441 4463 C2H2-TYPE.  
FT ZN\_FING 4471 4493 C2H2-TYPE.  
FT ZN\_FING 4501 4523 C2H2-TYPE.  
FT ZN\_FING 4531 4553 C2H2-TYPE.  
FT ZN\_FING 4561 4583 C2H2-TYPE.  
FT ZN\_FING 4591 4613 C2H2-TYPE.  
FT ZN\_FING 4621 4643 C2H2-TYPE.  
FT ZN\_FING 4651 4673 C2H2-TYPE.  
FT ZN\_FING 4681 4703 C2H2-TYPE.  
FT ZN\_FING 4711 4733 C2H2-TYPE.  
FT ZN\_FING 4741 4763 C2H2-TYPE.  
FT ZN\_FING 4771 4793 C2H2-TYPE.  
FT ZN\_FING 4801 4823 C2H2-TYPE.  
FT ZN\_FING 4831 4853 C2H2-TYPE.  
FT ZN\_FING 4861 4883 C2H2-TYPE.  
FT ZN\_FING 4891 4913 C2H2-TYPE.  
FT ZN\_FING 4921 4943 C2H2-TYPE.  
FT ZN\_FING 4951 4973 C2H2-TYPE.  
FT ZN\_FING 4981 5003 C2H2-TYPE.  
FT ZN\_FING 5011 5033 C2H2-TYPE.  
FT ZN\_FING 5041 5063 C2H2-TYPE.  
FT ZN\_FING 5071 5093 C2H2-TYPE.  
FT ZN\_FING 5101 5123 C2H2-TYPE.  
FT ZN\_FING 5131 5153 C2H2-TYPE.  
FT ZN\_FING 5161 5183 C2H2-TYPE.  
FT ZN\_FING 5191 5213 C2H2-TYPE.  
FT ZN\_FING 5221 5243 C2H2-TYPE.  
FT ZN\_FING 5251 5273 C2H2-TYPE.  
FT ZN\_FING 5281 5303 C2H2-TYPE.  
FT ZN\_FING 5311 5333 C2H2-TYPE.  
FT ZN\_FING 5341 5363 C2H2-TYPE.  
FT ZN\_FING 5371 5393 C2H2-TYPE.  
FT ZN\_FING 5401 5423 C2H2-TYPE.  
FT ZN\_FING 5431 5453 C2H2-TYPE.  
FT ZN\_FING 5461 5483 C2H2-TYPE.  
FT ZN\_FING 5491 5513 C2H2-TYPE.  
FT ZN\_FING 5521 5543 C2H2-TYPE.  
FT ZN\_FING 5551 5573 C2H2-TYPE.  
FT ZN\_FING 5581 5603 C2H2-TYPE.  
FT ZN\_FING 5611 5633 C2H2-TYPE.  
FT ZN\_FING 5641 5663 C2H2-TYPE.  
FT ZN\_FING 5671 5693 C2H2-TYPE.  
FT ZN\_FING 5701 5723 C2H2-TYPE.  
FT ZN\_FING 5731 5753 C2H2-TYPE.  
FT ZN\_FING 5761 5783 C2H2-TYPE.  
FT ZN\_FING 5791 5813 C2H2-TYPE.  
FT ZN\_FING 5821 5843 C2H2-TYPE.  
FT ZN\_FING 5851 5873 C2H2-TYPE.  
FT ZN\_FING 5881 5903 C2H2-TYPE.  
FT ZN\_FING 5911 5933 C2H2-TYPE.  
FT ZN\_FING 5941 5963 C2H2-TYPE.  
FT ZN\_FING 5971 5993 C2H2-TYPE.  
FT ZN\_FING 6001 6023 C2H2-TYPE.  
FT ZN\_FING 6031 6053 C2H2-TYPE.  
FT ZN\_FING 6061 6083 C2H2-TYPE.  
FT ZN\_FING 6091 6113 C2H2-TYPE.  
FT ZN\_FING 6121 6143 C2H2-TYPE.  
FT ZN\_FING 6151 6173 C2H2-TYPE.  
FT ZN\_FING 6181 6203 C2H2-TYPE.  
FT ZN\_FING 6211 6233 C2H2-TYPE.  
FT ZN\_FING 6241 6263 C2H2-TYPE.  
FT ZN\_FING 6271 6293 C2H2-TYPE.  
FT ZN\_FING 6301 6323 C2H2-TYPE.  
FT ZN\_FING 6331 6353 C2H2-TYPE.  
FT ZN\_FING 6361 6383 C2H2-TYPE.  
FT ZN\_FING 6391 6413 C2H2-TYPE.  
FT ZN\_FING 6421 6443 C2H2-TYPE.  
FT ZN\_FING 6451 6473 C2H2-TYPE.  
FT ZN\_FING 6481 6503 C2H2-TYPE.  
FT ZN\_FING 6511 6533 C2H2-TYPE.  
FT ZN\_FING 6541 6563 C2H2-TYPE.  
FT ZN\_FING 6571 6593 C2H2-TYPE.  
FT ZN\_FING 6601 6623 C2H2-TYPE.  
FT ZN\_FING 6631 6653 C2H2-TYPE.  
FT ZN\_FING 6661 6683 C2H2-TYPE.  
FT ZN\_FING 6691 6713 C2H2-TYPE.  
FT ZN\_FING 6721 6743 C2H2-TYPE.  
FT ZN\_FING 6751 6773 C2H2-TYPE.  
FT ZN\_FING 6781 6803 C2H2-TYPE.  
FT ZN\_FING 6811 6833 C2H2-TYPE.  
FT ZN\_FING 6841 6863 C2H2-TYPE.  
FT ZN\_FING 6871 6893 C2H2-TYPE.  
FT ZN\_FING 6901 6923 C2H2-TYPE.  
FT ZN\_FING 6931 6953 C2H2-TYPE.  
FT ZN\_FING 6961 6983 C2H2-TYPE.  
FT ZN\_FING 6991 7013 C2H2-TYPE.  
FT ZN\_FING 7021 7043 C2H2-TYPE.  
FT ZN\_FING 7051 7073 C2H2-TYPE.  
FT ZN\_FING 7081 7103 C2H2-TYPE.  
FT ZN\_FING 7111 7133 C2H2-TYPE.  
FT ZN\_FING 7141 7163 C2H2-TYPE.  
FT ZN\_FING 7171 7193 C2H2-TYPE.  
FT ZN\_FING 7201 7223 C2H2-TYPE.  
FT ZN\_FING 7231 7253 C2H2-TYPE.  
FT ZN\_FING 7261 7283 C2H2-TYPE.  
FT ZN\_FING 7291 7313 C2H2-TYPE.  
FT ZN\_FING 7321 7343 C2H2-TYPE.  
FT ZN\_FING 7351 7373 C2H2-TYPE.  
FT ZN\_FING 7381 7403 C2H2-TYPE.  
FT ZN\_FING 7411 7433 C2H2-TYPE.  
FT ZN\_FING 7441 7463 C2H2-TYPE.  
FT ZN\_FING 7471 7493 C2H2-TYPE.  
FT ZN\_FING 7501 7523 C2H2-TYPE.  
FT ZN\_FING 7531 7553 C2H2-TYPE.  
FT ZN\_FING 7561 7583 C2H2-TYPE.  
FT ZN\_FING 7591 7613 C2H2-TYPE.  
FT ZN\_FING 7621 7643 C2H2-TYPE.  
FT ZN\_FING 7651 7673 C2H2-TYPE.  
FT ZN\_FING 7681 7703 C2H2-TYPE.  
FT ZN\_FING 7711 7733 C2H2-TYPE.  
FT ZN\_FING 7741 7763 C2H2-TYPE.  
FT ZN\_FING 7771 7793 C2H2-TYPE.  
FT ZN\_FING 7801 7823 C2H2-TYPE.  
FT ZN\_FING 7831 7853 C2H2-TYPE.  
FT ZN\_FING 7861 7883 C2H2-TYPE.  
FT ZN\_FING 7891 7913 C2H2-TYPE.  
FT ZN\_FING 7921 7943 C2H2-TYPE.  
FT ZN\_FING 7951 7973 C2H2-TYPE.  
FT ZN\_FING 7981 8003 C2H2-TYPE.  
FT ZN\_FING 8011 8033 C2H2-TYPE.  
FT ZN\_FING 8041 8063 C2H2-TYPE.  
FT ZN\_FING 8071 8093 C2H2-TYPE.  
FT ZN\_FING 8101 8123 C2H2-TYPE.  
FT ZN\_FING 8131 8153 C2H2-TYPE.  
FT ZN\_FING 8161 8183 C2H2-TYPE.  
FT ZN\_FING 8191 8213 C2H2-TYPE.  
FT ZN\_FING 8221 8243 C2H2-TYPE.  
FT ZN\_FING 8251 8273 C2H2-TYPE.  
FT ZN\_FING 8281 8303 C2H2-TYPE.  
FT ZN\_FING 8311 8333 C2H2-TYPE.  
FT ZN\_FING 8341 8363 C2H2-TYPE.  
FT ZN\_FING 8371 8393 C2H2-TYPE.  
FT ZN\_FING 8401 8423 C2H2-TYPE.  
FT ZN\_FING 8431 8453 C2H2-TYPE.  
FT ZN\_FING 8461 8483 C2H2-TYPE.  
FT ZN\_FING 8491 8513 C2H2-TYPE.  
FT ZN\_FING 8521 8543 C2H2-TYPE.  
FT ZN\_FING 8551 8573 C2H2-TYPE.  
FT ZN\_FING 8581 8603 C2H2-TYPE.  
FT ZN\_FING 8611 8633 C2H2-TYPE.  
FT ZN\_FING 8641 8663 C2H2-TYPE.  
FT ZN\_FING 8671 8693 C2H2-TYPE.  
FT ZN\_FING 8701 8723 C2H2-TYPE.  
FT ZN\_FING 8731 8753 C2H2-TYPE.  
FT ZN\_FING 8761 8783 C2H2-TYPE.  
FT ZN\_FING 8791 8813 C2H2-TYPE.  
FT ZN\_FING 8821 8843 C2H2-TYPE.  
FT ZN\_FING 8851 8873 C2H2-TYPE.  
FT ZN\_FING 8881 8903 C2H2-TYPE.  
FT ZN\_FING 8911 8933 C2H2-TYPE.  
FT ZN\_FING 8941 8963 C2H2-TYPE.  
FT ZN\_FING 8971 8993 C2H2-TYPE.  
FT ZN\_FING 9001 9023 C2H2-TYPE.  
FT ZN\_FING 9031 9053 C2H2-TYPE.  
FT ZN\_FING 9061 9083 C2H2-TYPE.  
FT ZN\_FING 9091 9113 C2H2-TYPE.  
FT ZN\_FING 9121 9143 C2H2-TYPE.  
FT ZN\_FING 9151 9173 C2H2-TYPE.  
FT ZN\_FING 9181 9203 C2H2-TYPE.  
FT ZN\_FING 9211 9233 C2H2-TYPE.  
FT ZN\_FING 9241 9263 C2H2-TYPE.  
FT ZN\_FING 9271 9293 C2H2-TYPE.  
FT ZN\_FING 9301 9323 C2H2-TYPE.  
FT ZN\_FING 9331 9353 C2H2-TYPE.  
FT ZN\_FING 9361 9383 C2H2-TYPE.  
FT ZN\_FING 9391 9413 C2H2-TYPE.  
FT ZN\_FING 9421 9443 C2H2-TYPE.  
FT ZN\_FING 9451 9473 C2H2-TYPE.  
FT ZN\_FING 9481 9503 C2H2-TYPE.  
FT ZN\_FING 9511 9533 C2H2-TYPE.  
FT ZN\_FING 9541 9563 C2H2-TYPE.  
FT ZN\_FING 9571 9593 C2H2-TYPE.  
FT ZN\_FING 9601 9623 C2H2-TYPE.  
FT ZN\_FING 9631 9653 C2H2-TYPE.  
FT ZN\_FING 9661 9683 C2H2-TYPE.  
FT ZN\_FING 9691 9713 C2H2-TYPE.  
FT ZN\_FING 9721 9743 C2H2-TYPE.  
FT ZN\_FING 9751 9773 C2H2-TYPE.  
FT ZN\_FING 9781 9803 C2H2-TYPE.  
FT ZN\_FING 9811 9833 C2H2-TYPE.  
FT ZN\_FING 9841 9863 C2H2-TYPE.  
FT ZN\_FING 9871 9893 C2H2-TYPE.  
FT ZN\_FING 9901 9923 C2H2-TYPE.  
FT ZN\_FING 9931 9953 C2H2-TYPE.  
FT ZN\_FING 9961 9983 C2H2-TYPE.  
FT ZN\_FING 9991 10000 C2H2-TYPE.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X85327; CAA59677.1; -  
DR PIR: T06491; T06491.  
DR InterPro: IPR001362; Glyco\_hydro\_32.  
DR Pfam: PF00251; Glyco\_hydro\_32; 1.  
DR SMART: SM00640; Glyco\_32; 1.  
DR PROSITE: PS00609; GLYCOSYL\_HYDROL\_F32; 1.  
KW Hydrolase; Glycosylase; Glycoprotein; Signal.  
FT SIGNAL 1 22  
FT CHAIN 1 22  
FT SIGNAL 23 555  
FT CHAIN 23 555  
FT ACT\_SITE 140 140  
FT CARBOHYD 154 154  
FT CARBOHYD 181 181  
FT CARBOHYD 337 337  
SQ SEQUENCE 555 AA; 62655 MW; 7B54BE251CFD5D44 CRC64;  
-----  
Query Match  
Best Local Similarity 8.5%; Score 84; DB 1; Length 555;  
Matches 38; Conservative 22; Mismatches 52; Indels 50; Gaps 8;  
-----  
OY 2 PTWANEAKLSGYODAKMSLQODONKRYLAQNTOTTDVGGEFTCVTAEGKKILN 61  
DB 174 PTANKINSSSEFRDPTTS-----WLGR-----DGFWR-----VLGSKIDTKGM- 212  
OY 62 ATILKKNHLDLKE-----SHETITWAKADY-----TTENG---IKYE 98  
DB 213 -ATLKSKNFYDWVAKKPLHSAEGTGMECPDFVLDKLLRTGVDTSRNGDDVRRV 271  
OY 99 TCGTRTQTFEDVFSDYKNCVDIVFPKRGSDG-----DY 135  
DB 272 LKVSIDDTKRHDHYLGSDYDVVDFVPEKGFEDNPFVLRXDY 313  
-----  
RESULT 8  
RBA4\_ECHMU STANDARD: PRT: 223 AA.  
ID RBA4\_ECHMU  
AC 08G23: 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Probable Ras-related protein Rab-4A.  
OS Echinococcus multilocularis.  
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;  
OC Cyclophyllidae; Taeniidae; Echinococcus.  
OX NCBI\_Taxid=6211;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H-95;  
RX MEDLINE=20556267; PubMed=10973970;  
RA Brehm K., Jensen K., Frosch M.;  
RT "mRNA trans-splicing in the human parasitic cestode Echinococcus  
multilocularis.";  
RL J. Biol. Chem. 275:38311-38318(2000).  
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR  
CC TRAFFIC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

-----  
CC EMBL: AJ292375; CAC18549.1; -  
DR HSSP: P05713; 3RAB.  
DR InterPro: IPR003579; GTPase\_Rab.  
DR InterPro: IPR001806; Ras\_trnsmg.  
DR InterPro: IPR005225; Small\_GTP.  
DR Pfam: PF00071; ras; 1.  
DR PRINTS: PR00449; RASPRNSFRNNG.  
DR SMART: SM00175; RAB; 1.  
KW TIGRfams: TIGR00231; small\_GTP; 1.  
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
FT NP\_BIND 16 23  
FT NP\_BIND 64 68  
FT NP\_BIND 122 125  
FT NP\_BIND 38 46  
FT DOMAIN 221 221  
FT LIPID 223 223  
SQ SEQUENCE 223 AA; 24454 MW; 9EB6D6166E2089A4 CRC64;  
-----  
Query Match  
Best Local Similarity 8.2%; Score 81; DB 1; Length 223;  
Matches 22; Conservative 17; Mismatches 34; Indels 8; Gaps 3;  
-----  
OY 64 ILKKNKHLTLKESHETITWAKADYTTENGIRY-ETGTRTQTFEDVFVSDYKNCVY 122  
DB 118 ILVGNK--KDLRDTGQVTHWEANFPAQENGLOFETLSALGTENIDAF-----SCVRV 170  
OY 123 FVFKRGSDGDEGLWVSEDK 143  
DB 171 LLSKSGELGDRLLVGSNK 191  
-----  
RESULT 9  
C4BP\_BOVIN STANDARD: PRT: 610 AA.  
ID C4BP\_BOVIN  
AC Q28065;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE C4b-binding protein alpha chain precursor (c4bp).  
GN C4BP.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_Taxid=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95015909; PubMed=7930621;  
RA Hillarp A., Thern A., Dahlbaeck B.;  
RT "Bovine C4b binding protein. Molecular cloning of the alpha- and  
beta-chains provides structural background for lack of complex  
formation with protein S.";  
RL J. Immunol. 153:4190-4199(1994).  
CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT  
CC ACTIVATION. IT BINDS AS A COFACTOR TO C3B/C4B INACTIVATOR  
CC (C3BIN), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4B. IT  
CC ALSO ACCELERATES THE DEGRADATION OF THE C4B2A COMPLEX (C3  
CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2A. ALPHA  
CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.  
CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS  
CC (BY SIMILARITY).  
CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.  
CC -1- SIMILARITY: Contains 8 Sushi (SCR) domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

```

CC -----
DR EMBL: Z31693; CAAB3498.1;
DR PIR: I46001; I46001.
DR HSSP: P10998; IYVD.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; Sush1; 8.
DR SMART: SM00032; CCP; 8.
KW Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
FT SIGNAL 1 48
FT CHAIN 49 610
FT DOMAIN 49 108
FT DOMAIN 111 170
FT DOMAIN 173 235
FT DOMAIN 238 295
FT DOMAIN 298 363
FT DOMAIN 366 426
FT DOMAIN 428 484
FT DOMAIN 486 542
FT DOMAIN 50 95
FT DISULFID 80 107
FT DISULFID 112 153
FT DISULFID 139 169
FT DISULFID 174 217
FT DISULFID 203 234
FT DISULFID 239 281
FT DISULFID 267 294
FT DISULFID 299 350
FT DISULFID 334 362
FT DISULFID 366 390
FT DISULFID 367 425
FT DISULFID 429 471
FT DISULFID 457 483
FT DISULFID 487 528
FT DISULFID 514 541
FT DISULFID 549 561
FT DISULFID 561 561
FT CARBOHYD 66 66
FT CARBOHYD 221 221
FT CARBOHYD 525 525
FT CARBOHYD 602 602
SQ SEQUENCE 610 AA; 68886 MW; D806B270E8A06B58 CRC64;

Query Match 8.28; Score 81; DB 1; Length 610;
Best Local Similarity 23.7%; Pred. No. 13;
Matches 32; Conservative 14; Mismatches 47; Indels 42; Gaps 5;

QY 78 HETITWKAYD-----YTTENGIKYETGCTRTGFEDVVFVSDYKNCVIFVPRKRGSD 131
DB 315 HQTOTEGQVYDIFGLVLSYKCHFGYKREPTDPTVTCOSNLNEMSPYIECKEYCCPEPLUN 374
QY 132 EGYEL-----WSEDKID-----KIPDC---CKPTMA 156
DB 375 YGSITLHRRPSTSTHCTYISGDKISYCHSKYMEDALCTKHGTWSPRTPECRDCK-SP 433
QY 157 YFAQOQEKTVRNYT 171
DB 434 VLAHGKHVSKFEFT 448

RESULT 10
OPPF_MYCPN STANDARD; PRT; 851 AA.
AC P75551;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oligopeptide transport ATP-binding protein oppf.
GN OPF OR MPN218 OR MP613.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxId=2104;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE-97105885; PubMed-9948633.
RA Himmelreich R., Hilbert H., Plagens H., Pirkil E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION. PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR OLIGOPEPTIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC TRANSPORT SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000058; AAB96261.1;
DR PIR: S73939; S73939.
DR InterPro: IPR003593; AAA_ATPase.
DR Pfam: PF00005; ABC_tran; 1.
DR Prodom: PD000006; ABC_transporter; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 1.
KW Peptide transport; Transport; Membrane; ATP-binding;
FT NP_BIND 48 55
FT SEQUENCE 851 AA; 98611 MW; 6D3A3EC76B27DCC2 CRC64;

Query Match 8.28; Score 81; DB 1; Length 851;
Best Local Similarity 23.7%; Pred. No. 20;
Matches 44; Conservative 26; Mismatches 60; Indels 56; Gaps 8;

QY 3 TMANEAIGSYGDA-KKSLOO-DONKRYLAQNTOTTDGWSGEFFCVSTAEKIKKK- 59
DB 251 TIOQKAVSTQAFATWOKLTQTKONLKAIFAQ-----MAELONKPR 292
QY 60 -INATILYKXKHLTLKESHETTIVKAYDYTTENGIKYETGCTRTGFEDVVFVS--- 114
DB 293 IYLANML-----TTKNYIDSRQ--NTQLTDVFAFSYND 326
QY 115 -DYKNCVYFVPRKRGSDGDELYWSE--DKYDKIPDCKTTAVFAQOQEKTVRNY 170
DB 327 MVDKRRVLVLYSEYKALPEFYDNWIHONADRFDELTVNAVFEFLIDVYIALNRPANVE 386
QY 171 TDSSCK 176
DB 387 SDAKAE 392

RESULT 11
ATRX_HUMAN
ID ATRX_HUMAN STANDARD; PRT; 2492 AA.
AC P46100; P51068; Q15886; Q9H021; Q9NTS3;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
DE nuclear protein) (XNP) (znf-HX).
GN ATRX OR RAD54L OR XH2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANT SER-1860, AND

```

RP VARIANTS ATR-X.  
RA MEDLINE-97123494; PubMed-8968741;  
RA Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J.,  
RA Gibbons R.J.;  
RT "ATR-X encodes a novel member of the SNF2 family of proteins: mutations  
RT point to a common mechanism underlying the ATR-X syndrome.";  
RM Hum. Mol. Genet. 5:1899-1907(1996).  
RN [12]  
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).  
RX MEDLINE-97386582; PubMed-9244431.  
RA Villard L., Lissi A.-M., Cardoso C., Proud V., Chlaroni P.,  
RA Colleaux L., Schwartz C., Fontes M.;  
RT "Determination of the genomic structure of the XNP/ATR-X gene encoding  
RT a potential zinc finger helicase.";  
RN Genomics 43:149-155(1997).  
RN [13]  
RP SEQUENCE OF 860-2492 FROM N.A.  
RX MEDLINE-95179111; PubMed-7874112;  
RA Stayton C.L., Dabovic B., Gullisano M., Gecz J., Broccoli V.,  
RA Giovannuzzi S., Bossolasco M., Monaco L., Rastan S., Boncinelli E.,  
RA Bianchi M.E., Gonzalez G.G.;  
RT "Cloning and characterization of a new human Xq13 gene, encoding a  
RT putative helicase.";  
RN Hum. Mol. Genet. 3:1957-1964(1994).  
RN [14]  
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
RX MEDLINE-94214473; PubMed-8162050;  
RA Gecz J., Pollard H., Gonzalez G., Villard L., Stayton C.L.,  
RA Millaesau P., Khrestchatsky M., Fontes M.;  
RT "Cloning and expression of the murine homologue of a putative human  
RT X-linked nuclear protein gene closely linked to PKC1 in Xq13.3.";  
RN Hum. Mol. Genet. 3:39-44(1994).  
RN [15]  
RP SEQUENCE OF 2401-2492 FROM N.A., AND VARIANTS ATR-X.  
RX MEDLINE-95211835; PubMed-7697714;  
RA Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;  
RT "Mutations in a putative global transcriptional regulator cause X-  
RT linked mental retardation with alpha-thalassemia (ATR-X syndrome).";  
RN Cell 80:837-845(1995).  
RN [16]  
RP SEQUENCE OF 1375-2492 FROM N.A.  
RA Pearce A., Chapman J.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [17]  
RP E2H2 BINDING.  
RX MEDLINE-98167853; PubMed-9499421;  
RA Cardoso C., Timait S., Villard L., Khrestchatsky M., Fontes M.,  
RA Colleaux L.;  
RT "Specific interaction between the XNP/ATR-X gene product and the SET  
RT domain of the human E2H2 protein.";  
RN Hum. Mol. Genet. 7:679-684(1998).  
RN [18]  
RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC  
RP HETEROCHROMATIN.  
RX MEDLINE-20040663; PubMed-10570185;  
RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,  
RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,  
RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;  
RT "Localization of a putative transcriptional regulator (ATRX) at  
RT pericentromeric heterochromatin and the short arms of acrocentric  
RT chromosomes.";  
RN Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).  
RN [19]  
RP DISEASE.  
RX MEDLINE-20213147; PubMed-10751095;  
RA Villard L., Fontes M., Ades L.C., Gecz J.;  
RT "Identification of a mutation in the XNP/ATR-X gene in a family  
RT reported as Smith-Fineman-Hyers syndrome.";  
RN Am. J. Med. Genet. 91:83-85(2000).  
RN [10]  
RP VARIANTS ATR-X SER-1713.  
RX MEDLINE-97196774; PubMed-9043863;  
RA Villard L., Lacombe D., Fontes M.;

RT "A point mutation in the XNP gene, associated with an ATR-X phenotype  
RT without alpha-thalassemia.";  
RN Eur. J. Hum. Genet. 4:316-320(1996).  
RN [11]  
RP VARIANT JM GLN-2131.  
RX MEDLINE-96224392; PubMed-8630485;  
RA Villard L., Gecz J., Mattei J.-F., Fontes M., Saugier-Verber P.,  
RA Munnich A., Lyonnet S.;  
RT "XNP mutation in a large family with Juberger-Marsidi syndrome.";  
RN Nat. Genet. 12:359-360(1996).  
RN [12]  
RP VARIANTS ATR-X.  
RX MEDLINE-97467722; PubMed-9326931;  
RA Gibbons R.J., Bachoo S., Picketts D.J., Afifios S., Azenbauer B.,  
RA Begoffen J., Berry S.A., Dahl N., Fryer A., Keppeler K., Kurosawa K.,  
RA Levan M.L., Masuno M., Neri G., Pierpont M.E., Slaney S.F.,  
RA Higgs D.R.;  
RT "Mutations in transcriptional regulator ATRX establish the functional  
RT significance of a PHD-like domain.";  
RN Nat. Genet. 17:146-148(1997).  
RN [13]  
RP VARIANT ATR-X LEU-246.  
RX MEDLINE-20123062; PubMed-10660327;  
RA Fichera M., Romano C., Castiglia L., Falla P., Ruberto C., Amata S.,  
RA Greco D., Cardoso C., Fontes M., Ragusa A.;  
RT "New mutations in XNP/ATR-X gene: a further contribution to  
RT genotype/phenotype relationship in ATR-X syndrome.";  
RN Hum. Mutat. 12:214-214(1998).  
RN [14]  
RP VARIANT SHS LYS-1742.  
RX MEDLINE-99347960; PubMed-10417298;  
RA Lissi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,  
RA Prieto F., Fontes M., Martinez F.;  
RT "Mutation of the XNP/ATR-X gene in a family with severe mental  
RT retardation, spastic paraplegia and skewed pattern of X inactivation:  
RT demonstration that the mutation is involved in the inactivation  
RT bias.";  
RN Am. J. Hum. Genet. 65:558-562(1999).  
RN [15]  
RP VARIANT CMS THR-2050.  
RX MEDLINE-99326061; PubMed-10398237;  
RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,  
RA Curtis M.;  
RT "Carpenter-Waziri syndrome results from a mutation in XNP.";  
RN Am. J. Med. Genet. 85:249-251(1999).  
RN [16]  
RP VARIANTS ATR-X GIU-175; 178-VAL-LYS-198 DEL; SER-190; PRO-219;  
RX LEU-246 AND CYS-249.  
RX MEDLINE-99219535; PubMed-10204841;  
RA Villard L., Bonlino M.-C., Abidi F., Ragusa A., Beloune J.,  
RA Lissi A.-M., Seaver L., Bonnefont J.-P., Romano C., Fichera M.,  
RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;  
RT "Evaluation of a mutation screening strategy for sporadic cases of  
RT ATR-X syndrome.";  
RN J. Med. Genet. 36:183-186(1999).  
RN [17]  
RP VARIANTS ATR-X SER-179; LEU-190; ILE-194; CYS-246; PHE-1552; SER-1645  
RX AND CYS-1847.  
RX MEDLINE-20451413; PubMed-10995512;  
RA Wada T., Kubota T., Fukushima Y., Saitoh S.;  
RT "Molecular genetic study of Japanese patients with X-linked alpha-  
RT thalassemia/mental retardation syndrome (ATR-X).";  
RN Am. J. Med. Genet. 94:242-248(2000).  
RN [18]  
RP FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES  
RP GENE EXPRESSION BY AFFECTING CHROMATIN. MAY BE INVOLVED IN BRAIN  
RP DEVELOPMENT AND FACIAL MORPHOGENESIS.  
CC -1- SUBUNIT: PROBABLY BINDS E2H2. BINDS ANNEXIN V IN A CALCIUM AND  
CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSERINE-DEPENDENT MANNER (BY  
CC similarity).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC  
CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY  
CC INTERACTING WITH HP1.  
CC -1- ALTERNATIVE PRODUCTS:



CC Event-Alternative splicing; Named isoforms=5;  
CC Name=4;  
CC IsoId=P46100-1; Sequence-Displayed;  
CC Name=1;  
CC IsoId=P46100-2; Sequence-VSP\_000575;  
CC Name=2;  
CC IsoId=P46100-3; Sequence-VSP\_000574;  
CC Name=3;  
CC IsoId=P46100-4; Sequence-VSP\_000576;  
CC Name=5;  
CC IsoId=P46100-5; Sequence-VSP\_000574, VSP\_000576;  
CC TISSUE SPECIFICITY: Ubiquitous;  
CC -1- DISEASE: Defects in ATRX are the cause of X-linked alpha-thalassemia/mental retardation syndrome (ATR-X) [MIM:301040]. ATR-X is an X-linked disorder comprising severe psychomotor retardation, facial dysmorphism, urogenital abnormalities, and alpha-thalassemia. An essential phenotypic trait are hemoglobin H erythrocyte inclusions.  
CC -1- DISEASE: Defects in ATRX are the cause of Sutherland-Haan X-linked mental retardation syndrome (SHS) [MIM:309470]. It is characterized by severe mental retardation with spastic paraplegia, microcephaly, short stature and cryptorchidism.  
CC -1- DISEASE: Defects in ATRX are a cause of Smith-Fineman-Wyers syndrome (SFM) [MIM:309580]. Clinical features include severe mental retardation, microcephaly, growth failure, facial anomalies and bilateral cryptorchidism. Due to the clinical overlap with ATR-X syndrome, some patients originally diagnosed as having SFM, might be affected by a variant of ATR-X syndrome which lack hemoglobin h inclusions.  
CC -1- DISEASE: Defects in ATRX are the cause of Carpenter-Wariri syndrome (CMS), an X-linked recessive condition characterized by moderate mental retardation, short stature, brachydactyly with excessive skin creases, and widening of the knuckles.  
CC -1- DISEASE: Defects in ATRX are the cause of Juber-Marsidi syndrome (JM) [MIM:309590]. JM is a rare X-linked recessive disease characterized by severe mental retardation, growth failure, sensorineural deafness, microgenitalism and early death.  
CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.  
CC -----  
Query Match 8.2%; Score 81; DB 1; Length 2492;  
Best Local Similarity 26.4%; Pred. No. 71;  
Matches 42; Conservative 21; Mismatches 56; Indels 40; Gaps 8;  
OY 32 QATQTGTV---WGEFTCVSYTAERIGKKKLNATLLYKNKHLTDKESHETTIVKAYD 88  
DB 913 QASASTDGDYDKLSGKEQSFSLSEVRKVAETK-----EKSRL-----KTKCKKVD 959  
OY 89 YTTENGKIKYETGCTRTQTFEDVFVSDYKNCQVIFPKKGGSP-----EDDYELWS 140  
DB 960 GLSADIKFLKQKQSDSTSD-----DKOSKKGTEKKKPSDFKKRVIKMEQOYE--SS 1012  
OY 141 EDKIDKIP---DCCKFTMAVFAOQOEKTVNRYNTDSSCK 176  
DB 1013 SDGTEKLPREKICHPKGI-----KQIKNGTTDDEKK 1045  
RESULT 12  
CPAH MOUSE STANDARD; PRT; 1234 AA.  
AC P06909;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Complement factor H precursor (Protein beta-1-H).  
GN HFI OR CFH  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=86233353; PubMed=2940596;  
RA Kristensen T., Tack B.F.;  
RT "Murine protein H is comprised of 20 repeating units, 61 amino acids in length."  
RL Proc. Natl. Acad. Sci. U.S.A. 83:3963-3967(1986).  
RN [2]  
RP SEQUENCE OF 1-19 FROM N.A.  
RC STRAIN=BA1B/C;  
RX MEDLINE=90148935; PubMed=2533512;  
RA Munoz-Canoves P., Tack B.F., Vik D.P.;  
RT "Analysis of complement factor H mRNA expression: dexamethasone and IFN-gamma increase the level of H in L cells."  
RL Biochemistry 28:9891-9897(1989).  
RN [3]  
RP SEQUENCE OF 1-18 FROM N.A.  
RX MEDLINE=90111033; PubMed=2136885;  
RA Natsume-Sakai S., Nonaka M., Nonaka M., Harada Y.N., Shreffler D.C., Moriaki K.;  
RT "Demonstration of an unusual allelic variation of mouse factor H by the complete cDNA sequence of the H.2 allotype."  
RL J. Immunol. 144:358-362(1990).  
CC -1- FUNCTION: Factor H functions as a cofactor in the inactivation of C3b by factor I and also increases the rate of dissociation of the convertase (C3 convertase) and the (C3b)NBB complex (C5 convertase) in the alternative complement pathway (by similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.  
CC -1- POLYMORPHISM: TWO CODOMINANT ALLELES OF FACTOR H ARE PRESENT IN MICE.  
CC -----  
CC -1- SIMILARITY: Contains 20 Sushi (SCR) domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M12660; AAA37759.1; -;  
DR EMBL; J02891; AAA37795.1; -;  
DR EMBL; M31979; AAA37762.1; -;  
DR PIR; A26154; NEMSH.  
DR HSSP; P08603; IHFI.  
DR MGD; MGI:88385; CfH.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00084; sushi; 20.  
DR SMART; SM00032; CCP; 20.  
KW Complement alternate pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal.  
FT 1 18 BY SIMILARITY.  
FT CHAIN 19 1234 COMPLEMENT FACTOR H.  
FT DOMAIN 20 81 SUSHI 1.  
FT DOMAIN 84 142 SUSHI 2.  
FT DOMAIN 145 206 SUSHI 3.  
FT DOMAIN 209 263 SUSHI 4.  
FT DOMAIN 266 321 SUSHI 5.  
FT DOMAIN 324 386 SUSHI 6.  
FT DOMAIN 388 443 SUSHI 7.  
FT DOMAIN 447 506 SUSHI 8.  
FT DOMAIN 508 565 SUSHI 9.  
FT DOMAIN 568 623 SUSHI 10.  
FT DOMAIN 628 684 SUSHI 11.  
FT DOMAIN 689 744 SUSHI 12.  
FT DOMAIN 751 803 SUSHI 13.  
FT DOMAIN 807 862 SUSHI 14.  
FT DOMAIN 866 932 SUSHI 15.  
FT DOMAIN 935 990 SUSHI 16.  
FT DOMAIN 993 1049 SUSHI 17.  
FT DOMAIN 1052 1108 SUSHI 18.  
FT DOMAIN 1113 1169 SUSHI 19.  
FT DOMAIN 1171 1234 SUSHI 20.  
FT DISULFID 21 66 BY SIMILARITY.



```

FT DISULFID 52 80 BY SIMILARITY.
FT DISULFID 85 129 BY SIMILARITY.
FT DISULFID 114 141 BY SIMILARITY.
FT DISULFID 146 192 BY SIMILARITY.
FT DISULFID 178 205 BY SIMILARITY.
FT DISULFID 210 251 BY SIMILARITY.
FT DISULFID 237 262 BY SIMILARITY.
FT DISULFID 267 309 BY SIMILARITY.
FT DISULFID 294 320 BY SIMILARITY.
FT DISULFID 325 374 BY SIMILARITY.
FT DISULFID 357 385 BY SIMILARITY.
FT DISULFID 389 431 BY SIMILARITY.
FT DISULFID 416 442 BY SIMILARITY.
FT DISULFID 448 494 BY SIMILARITY.
FT DISULFID 477 505 BY SIMILARITY.
FT DISULFID 509 553 BY SIMILARITY.
FT DISULFID 536 564 BY SIMILARITY.
FT DISULFID 569 610 BY SIMILARITY.
FT DISULFID 597 622 BY SIMILARITY.
FT DISULFID 629 672 BY SIMILARITY.
FT DISULFID 658 683 BY SIMILARITY.
FT DISULFID 690 732 BY SIMILARITY.
FT DISULFID 718 743 BY SIMILARITY.
FT DISULFID 752 791 BY SIMILARITY.
FT DISULFID 780 802 BY SIMILARITY.
FT DISULFID 808 850 BY SIMILARITY.
FT DISULFID 836 861 BY SIMILARITY.
FT DISULFID 867 920 BY SIMILARITY.
FT DISULFID 906 931 BY SIMILARITY.
FT DISULFID 936 978 BY SIMILARITY.
FT DISULFID 964 989 BY SIMILARITY.
FT DISULFID 1023 1048 BY SIMILARITY.
FT DISULFID 1053 1096 BY SIMILARITY.
FT DISULFID 1082 1107 BY SIMILARITY.
FT DISULFID 1114 1157 BY SIMILARITY.
FT DISULFID 1143 1168 BY SIMILARITY.
FT DISULFID 1172 1223 BY SIMILARITY.
FT DISULFID 1206 1233 BY SIMILARITY.
FT CARBOHYD 676 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 801 801 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1225 1225 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1234 AA; 139082 MW; C5AC02F341B957F7 CRC64;

```

Query Match 8.1%; Score 80.5; DB 1; Length 1234;  
 Best Local Similarity 29.8%; Pred. No. 34;  
 Matches 25; Conservative 13; Mismatches 19; Indels 27; Gaps 5;

```

OY 24 QNKRYYLAQATQT---TDGVGGEFTCVSTAEEKIGKKKLNATLLYKKHLTDLKESHEI 80
DB 1144 QOCRRYYLKGKRTTCTGNGKSEPTCTC-----LHACVYIPEN-----IMESHNI 1184
OY 81 ITVM-----KAYDYTTEN-----GIRY 97
DB 1185 ILKWRHTEKITSHSGEDIEFGCKY 1208

```

RESULT 13  
 PCN3\_SULTO STANDARD: PRT: 246 AA.  
 AC 0973F5  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE DNA polymerase sliding clamp C (Proliferating cell nuclear antigen homolog C) (PCNA C).  
 GN PCNC OR S70944.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

```

CC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankel A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Koshida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermocacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
CC -1- FUNCTION: Sliding clamp subunit. Responsible for tethering the
CC catalytic subunit of DNA polymerase to DNA during high-speed
CC replication (By similarity).
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PCNA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF000984; BAB65958.1; -
DR HAMAP; MF_00317; -; 1.
DR InterPro; IPR000730; Pr_cel_nuc_antlg.
DR Pfam; PF00705; PCNA_1.
DR PRINTS; PR00339; PCNACTCLIN.
DR ProDom; PD002673; Pr_cel_nuc_antlg; 1.
DR PROSITE; PS01251; PCNA_1; FALSE_NEG.
DR PROSITE; PS00293; PCNA_2; FALSE_NEG.
DR DNA-binding; DNA replication; Complete proteome.
SQ SEQUENCE 246 AA; 27436 MW; BA370A5D27B4E7F CRC64;

```

Query Match 8.1%; Score 80; DB 1; Length 246;  
 Best Local Similarity 26.2%; Pred. No. 5.6;  
 Matches 33; Conservative 17; Mismatches 38; Indels 38; Gaps 5;

```

OY 9 KLSGYDAWKS-----LQDDQNKRRYYLAQATQTDGVGGEFTCVSTAEEKIGKKKL 60
DB 70 KLEDFDVLKVTYKNDLSLYLETDENQIKV-----TLDGVYERFTFPPIVASEIETPNL 124
OY 61 NATLLYKKHL-----TDLKESHEI-----TWKKA-----YDYTTENGI 95
DB 125 NLEFPFKAKALTVFTFDIIDLIEDIGDSITFKMGCKLYLSANSDMGSSITELSTENG 184
OY 96 KYETPG 101
DB 185 LLESEG 190

```

RESULT 14  
 PEPL\_YEAST STANDARD: PRT: 1579 AA.  
 AC P32319;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Vacuolar protein sorting/targeting protein PEPL precursor (Vacuolar carboxypeptidase sorting/receptor VPS10) (Carboxypeptidase Y receptor) (CPY receptor).  
 GN PEPL OR VPS10 OR YBL017C OR YBL0302 OR YBL03.22.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE-9307614; PubMed-1441754;  
 RA van Dyck L., Purnelle B., Skala J., Goffeau A.;  
 RT "An 11.4 kb DNA segment on the left arm of yeast chromosome II  
 RT carries the carboxypeptidase Y sorting gene PEPI, as well as ACH1,  
 RT FUS3 and a putative ARS.";  
 RL Yeast 8:769-776(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94243924; PubMed-8187177;  
 RA Marcusson E.G., Horadzovsky B.F., Cereghino J.L., Gharkhanian E.,  
 RA Ent S.D.;  
 RT "The sorting receptor for yeast vacuolar carboxypeptidase Y is  
 RT encoded by the VPS10 gene.";  
 RL Cell 77:579-586(1994).  
 CC -1- FUNCTION: INVOLVED IN THE INTRACELLULAR SORTING AND DELIVERY OF  
 CC SOLUBLE VACUOLAR PROTEINS, SEEMS TO SPECIFICALLY INTERACT WITH  
 CC CARBOXYPEPTIDASE Y (CPY), MAY EXECUTE MULTIPLE ROUNDS OF SORTING  
 CC BY CYCLING BETWEEN THE LATE GOLGI AND A PREVACUOLAR ENDOSOME-LIKE  
 CC COMPARTMENT. BINDS THE GOLGI-MODIFIED P2 FORM OF CPY, AND THIS  
 CC INTERACTION IS DEPENDENT ON THE PRESENCE OF AN INTACT CPY VACUOLAR  
 CC PROTEIN SORTING SIGNAL.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. LATE GOLGI  
 CC COMPARTMENT.  
 CC -1- DOMAIN: THE LUMENAL DOMAIN CONTAINS TWO REGIONS OF APPROXIMATELY  
 CC 650 AA THAT EXHIBIT 20% IDENTITY. THE CYTOPLASMIC DOMAIN MAY SERVE  
 CC AS A GOLGI RETENTION/RECYCLING SIGNAL.  
 CC -1- SIMILARITY: BELONGS TO THE PEPI FAMILY OF MEMBRANE GLYCOPROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: X68577; CAA48568.1; -;  
 DR EMBL: Z35778; CAA84836.1; -;  
 DR EMBL: U07621; AAA1831.1; -;  
 DR PIR: S25329; S25329.  
 DR SCD: S0000113; PEPI.  
 DR GO: GO:0006623; P:protein-vacuolar targeting; IMP.  
 DR GO: GO:0007034; P:vacuolar transport; IMP.  
 DR InterPro: IPR002860; GH\_BNR.  
 DR Pfam: PFO2012; BNR. 13.  
 DR SMART: SM00602; VPS10. 2.  
 KW Transmembrane; Glycoprotein; Receptor; Signal; ATP-binding; Repeat.  
 FT SIGNAL 1 21  
 FT CHAIN 22 1579  
 FT FT  
 FT DOMAIN 22 1397  
 FT TRANSMEM 1398 1414  
 FT DOMAIN 1415 1579  
 FT NP\_BIND 180 187  
 FT CARBOHYD 96 96  
 FT CARBOHYD 170 170  
 FT CARBOHYD 447 447  
 FT CARBOHYD 793 793  
 FT CARBOHYD 1010 1010  
 FT CARBOHYD 1303 1303  
 FT CARBOHYD 19 19  
 FT CONFLICT 38 38  
 FT CONFLICT 54 54  
 FT CONFLICT 74 74  
 FT CONFLICT 94 94  
 FT CONFLICT 119 126  
 FT CONFLICT 152 153  
 FT CONFLICT 165 167  
 FT CONFLICT 202 202  
 FT CONFLICT 214 214  
 S -> G (IN REF. 2).  
 S -> G (IN REF. 2).

FT CONFLICT 283 283 L -> I (IN REF. 2).  
 FT CONFLICT 378 378 F -> L (IN REF. 2).  
 FT CONFLICT 410 410 S -> N (IN REF. 2).  
 FT CONFLICT 426 426 K -> V (IN REF. 2).  
 FT CONFLICT 774 774 M -> R (IN REF. 2).  
 FT CONFLICT 1089 1089 S -> F (IN REF. 2).  
 FT CONFLICT 1266 1266 G -> S (IN REF. 2).  
 FT CONFLICT 1476 1476 A -> V (IN REF. 2).  
 FT CONFLICT 1557 1557 T -> A (IN REF. 2).  
 SQ SEQUENCE 1579 AA; 177776 MW; C778A0400612ECFA CRC64;  
 Query Match 8.1%; Score 80; DB 1; Length 1579;  
 Best Local Similarity 27.8%; Pred. No. 50;  
 Matches 35; Conservative 12; Mismatches 37; Indels 42; Gaps 7;  
 QY 84 WKADYTEGNIK-EMOGRTRQFEDEVFV-----SDYKNCVIFVPEK 127  
 Db 607 WNLAEKCVAGVKKKIRKRODACLKVKFEDIQLETACDKCTEADYE-CAFEYRDA 665  
 QY 128 RGSDEGDEYELWSEDKIDK-----IP-----DCKFTMAVFAQOQERTVRYVYDS 173  
 Db 666 TGKCVDPYNIIVLSDVCDKTKKTVPVKPLQLVKGDKCKPM-----TVKSY--DI 714  
 QY 174 SCKRPAP 179  
 Db 715 SCRGVP 720  
 RESULT 15  
 ID POLG\_EC01F STANDARD; PRT; 2184 AA.  
 AC 091734; Q66795; Q9Y1D6;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [Contigs: Coat protein VP4 (P1A); Coat protein VP2  
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A  
 DE (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein  
 DE P3A; Genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28)  
 DE (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
 OS Echovirus 1 (strain Farout / ATCC VR-1038).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=103908;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bergelson J.M.;  
 RT "Receptor interactions, infectious cDNA, and nucleotide sequences of  
 RT echovirus 1/8." to the EMBL/GenBank/DBJ databases.  
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 69-330 FROM N.A.  
 RX MEDLINE-96203959; PubMed-8627260;  
 RA Huttunen P., Santti J., Pulli T., Hyypiae T.;  
 RT "The major echovirus group is genetically coherent and related to  
 RT coxsackie B viruses.";  
 RL J. Gen. Virol. 77:715-725(1996).  
 RN [3]  
 RP SEQUENCE OF 570-853 FROM N.A.  
 RX MEDLINE-9918973; PubMed-9971773;  
 RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;  
 RT "Molecular evolution of the human enteroviruses: correlation of  
 RT serotype with VP1 sequence and application to picornavirus  
 RT classification.";  
 RL J. Virol. 73:1941-1948(1999).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (3.55 ANGSTROMS) OF 2-850.  
 RX MEDLINE-99192672; PubMed-10089503;  
 RA Filman D.J., Wien M.W., Cunningham J.A., Bergelson J.M., Hogle J.M.;  
 RT "Structure determination of echovirus 1.";  
 RL Acta Crystallogr. D 54:1261-1272(1998).  
 CC -1- FUNCTION: P2A AND THE P3C POLYPEPTIDES ARE PROTEASES THAT CLEAVE  
 CC AT CERTAIN Q/G SITES IN THE POLYPROTEIN. THEY ARE CYSTEINE

CC PROTEASES.  
CC -1- FUNCTION: It is thought that the p2c protein attaches to vesicular  
CC membranes and is associated with viral RNA synthesis.  
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-gly bond in the  
CC poliovirus polypeptide. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Tyr-I-gly bond in the  
CC picornavirus polypeptide. In other picornavirus reactions Glu may  
CC be substituted for Gln, and Ser or Thr for Gly.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC [RNA](N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC. VP1/VP2 IS  
CC CATALYZED BY P2A. ALL OTHER CLEAVAGES ARE CATALYZED BY P3C.  
CC -1- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF029859; AACG3944.2; -  
CC EMBL: X89531; CA61710.1; -  
CC EMBL: AF081314; AADI7718.1; -  
CC PDB: 1EV1, 27-JAN-99.  
CC InterPro: IPR003593; AAA\_ATPase.  
CC InterPro: IPR004004; Calic\_pol\_hel.  
CC InterPro: IPR000199; Cys\_protease-3c.  
CC InterPro: IPR003138; Pico\_P1A.  
CC InterPro: IPR000081; Pico\_P2A.  
CC InterPro: IPR002527; Pico\_P2B.  
CC InterPro: IPR001676; RHV.  
CC InterPro: IPR000605; RNA\_helicase.  
CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro: IPR001205; RNA\_pol\_P3D.  
CC InterPro: IPR007094; RNA\_pol\_PSVir.  
CC Pfam: PF00548; Cys-Protease-3c; 1.  
CC Pfam: PF02226; Pico\_P1A; 1.  
CC Pfam: PF00947; Pico\_P2A; 1.  
CC Pfam: PF01552; Pico\_P2B; 1.  
CC Pfam: PF00073; TNV; 3.  
CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
CC Pfam: PF00910; RNA\_helicase; 1.  
CC PRINTS: PR00918; CALICVIRUSNS.  
CC ProDom: PD001125; Cys\_protease-3c; 1.  
CC ProDom: PD001306; Pico\_P2A; 1.  
CC ProDom: PD001274; Pico\_P2B; 1.  
CC SMART: SM00382; AAA; 1.  
CC Polyprotein; Coat protein; Core protein; Transferase; Myristate;  
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
CC RNA  
CC CHAIN 2 69 COAT PROTEIN VP4.  
CC CHAIN 70 330 COAT PROTEIN VP2.  
CC CHAIN 331 569 COAT PROTEIN VP3.  
CC CHAIN 570 850 COAT PROTEIN VP1.  
CC CHAIN 851 1000 PICORNAIN 2A.  
CC CHAIN 1001 1099 CORE PROTEIN P2B.  
CC CHAIN 1100 1428 CORE PROTEIN P2C.  
CC CHAIN 1429 1517 CORE PROTEIN P3A.  
CC CHAIN 1518 1539 GENOME-LINKED PROTEIN VPG.  
CC CHAIN 1540 1722 PICORNAIN 3C.  
CC CHAIN 1723 2184 RNA-DIRECTED RNA POLYMERASE.  
CC LIPID 2 MYRISTATE.  
CC ACT\_SITE 1686 1686 PROTEASE 3C (POTENTIAL).  
CC ACT\_SITE 1700 1700 PROTEASE 3C (POTENTIAL).  
CC CONFLICT 69 69 N -> M (IN REF. 2).  
CC CONFLICT 238 238 A -> R (IN REF. 2).

FT CONFLICT 611 611 A -> V (IN REF. 3).  
FT CONFLICT 645 645 Y -> F (IN REF. 3).  
FT CONFLICT 724 724 V -> I (IN REF. 3).  
FT CONFLICT 774 774 A -> T (IN REF. 3).  
FT STRAND 4 7  
FT STRAND 26 29  
FT HELIX 36 38  
FT STRAND 45 45  
FT TURN 50 50  
FT HELIX 51 54  
FT STRAND 57 57  
FT TURN 63 64  
FT TURN 80 81  
FT STRAND 83 86  
FT STRAND 91 94  
FT STRAND 101 102  
FT HELIX 103 105  
FT TURN 113 115  
FT STRAND 123 123  
FT HELIX 126 128  
FT TURN 129 129  
FT STRAND 133 134  
FT STRAND 138 140  
FT TURN 142 143  
FT STRAND 147 151  
FT TURN 152 152  
FT HELIX 153 157  
FT STRAND 159 160  
FT STRAND 168 180  
FT TURN 185 186  
FT STRAND 188 197  
FT TURN 198 198  
FT STRAND 203 203  
FT STRAND 206 207  
FT HELIX 212 214  
FT TURN 215 215  
FT STRAND 217 217  
FT TURN 218 219  
FT STRAND 220 220  
FT STRAND 222 223  
FT STRAND 223 223  
FT TURN 232 233  
FT STRAND 235 235  
FT STRAND 238 238  
FT STRAND 239 241  
FT HELIX 242 245  
FT STRAND 246 246  
FT HELIX 248 253  
FT STRAND 256 260  
FT TURN 261 263  
FT STRAND 266 271  
FT STRAND 280 280  
FT TURN 282 284  
FT STRAND 285 285  
FT STRAND 288 299  
FT TURN 302 303  
FT STRAND 308 324  
FT TURN 338 341  
FT TURN 345 346  
FT STRAND 353 353  
FT TURN 356 357  
FT STRAND 369 370  
FT STRAND 372 372  
FT TURN 373 373  
FT HELIX 374 377  
FT TURN 378 378  
FT STRAND 381 382  
FT HELIX 390 393  
FT TURN 394 394  
FT HELIX 397 399  
FT STRAND 400 405  
FT TURN 406 406  
FT TURN 410 410

FT	STRAND	412	417
FT	TURN	420	422
FT	HELIX	424	427
FT	TURN	428	428
FT	HELIX	430	435
FT	TURN	436	437
FT	STRAND	438	442
FT	STRAND	445	452
FT	TURN	455	456
FT	STRAND	458	458
FT	STRAND	460	466

Query Match	8.08;	Score 79.5;	DB 1;	Length 2184;
Best Local Similarity	32.14;	Pred No 87;		

```

QY      5  ANEKILISYDQAMKSLOODOKRRYYLQAOTOTTOGVNGEERT-----CVSTJARIKIGCK  59
Db      1068  ATTLALIGCTSSPFWMLKOKVSOYYGIPRAEKQNGG-WIKKFTMTNACKGH--EWIAIKI  1124

QY      60  LNAATILYKNKHULTJLKESEHERITVWAKAVDYTTENGIRYE-----TQGTRTQTFEDVFESD  115
Db      1125  QKFTIEMUKVKILPEVKEKEHEFLNMLKQDLPLESGIATIEGASPSQSGDQELFSNVQYFAH  1184

QY      116  YKNCDDVIFVPERKESDSDGDEYELWVSEDKIDKIPDCKEFTMAVFAQQQ-----EKTVRN-V  169
Db      1185  Y-----CRKYAPLPYAAEAKRVFSLERKMSNYI  1211

QY      170  YTDSSCKPAP  179
Db      1212  QFKSKRIEP  1221

```

```
Search completed: August 1, 2003, 12:15:53
Job time : 13.0933 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 : Search time 13.52 Seconds

(without alignments)  
1294.578 Million cell updates/sec

Title: US-10-087-195-8

Sequence: 1 NPTWANEAKLGSYQDAWKS.....EKTVRNVYTDSCKPAPQN 182

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR\_76:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.5	11.5	330	2 S70127	hypothetical prote
2	87	8.8	483	2 ARI609	two-component sens
3	86	8.7	529	2 S18453	variant surface gl
4	85.5	8.6	1970	2 T03284	myoblast city prot
5	85	8.6	1959	2 AG1085	hypothetical prote
6	84	8.5	555	2 T06491	beta-fructofuranos
7	84	8.5	582	2 T06380	beta-fructofuranos
8	83.5	8.4	1078	2 T00494	RNA-directed RNA p
9	82.5	8.3	306	2 T12821	hypothetical prote
10	82.5	8.3	1802	2 G71616	hypothetical prote
11	81	8.2	610	1 I46001	C4d-binding protei
12	81	8.2	656	2 S30484	pol polypeptide
13	81	8.2	851	2 S73939	oligopeptide trans
14	81	8.2	851	2 S73939	oligopeptide trans
15	81	8.2	1641	2 I18614	helicase II - huma
16	80.5	8.1	721	2 ARI254	penicillin-binding
17	80.5	8.1	1234	1 NBMSH	complement factor
18	80	8.1	641	2 PH1919	FL-160-4 protein -
19	80	8.1	1579	2 S25329	carboxypeptidase y
20	79	8.0	338	2 S75154	hypothetical prote
21	79	8.0	345	2 C89882	conserved hypotet
22	79	8.0	375	2 AC1350	N-acetylmuramoyl-L
23	79	8.0	405	2 G81184	late embryogenesis
24	79	8.0	479	2 T47561	glutaminyl-tRNA sy
25	79	8.0	580	2 G82693	glutaminyl-tRNA sy
26	79	8.0	1405	2 T04426	hypothetical prote
27	79	8.0	1607	2 T03022	MAP kinase kinase
28	78.5	7.9	212	2 F64236	hypothetical prote
29	78.5	7.9	320	2 S51051	glutamate COA-tr

30	77.5	7.8	306	2 C64015	hypothetical prote
31	77.5	7.8	423	2 AB1142	N-carbamyl-L-amino
32	77.5	7.8	769	2 P89870	serine proteinase
33	77.5	7.8	1156	2 B70356	chromosome assembl
34	77	7.8	231	2 T41566	translin associate
35	77	7.8	334	2 T47983	NM-1-like protein -
36	77	7.8	421	2 C70038	maltose/maltodextr
37	77	7.8	484	2 A64251	glutamate-tRNA 119
38	77	7.8	656	2 S30483	pol polypeptide
39	77	7.8	719	2 S44237	trypanastigote sur
40	77	7.8	762	2 JC7114	N,N-dimethylformam
41	77	7.8	1108	2 T38673	probable transcrip
42	77	7.8	2034	2 T22147	hypothetical prote
43	77	7.8	2245	2 T18278	myosin heavy chain
44	76.5	7.7	414	2 ARI500	N-carbamyl-L-amino
45	76.5	7.7	434	2 C82885	membrane nuclease

#### ALIGNMENTS

##### RESULT 1

S70127 hypothetical protein YDR267c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D9954.11

C:Species: Saccharomyces cerevisiae

C:Date: 24-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 19-Apr-2002

C:Accession: S70127

R:Le, T.

submitted to the EMBL Data Library, May 1996

A:Description: The sequence of S. cerevisiae cosmid 9954.

A:Reference number: S70124

A:Accession: S70127

A:Molecule type: DNA

A:Residues: 1-330 <LET>

A:Cross-references: EMBL:U51030; NID:G1332633; PID:G1230640; GSPDB:GN00004; MIPS:YDR2

C:Genetics:

A:Gene: MIPS:YDR267c

A:Cross-references: SGD:S0002675

A:Map position: 4R

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

F:54-87/Domain: WD repeat homology <WD1>

F:103-136/Domain: WD repeat homology <WD2>

F:149-182/Domain: WD repeat homology <WD3>

Query Match

Best Local Similarity 11.5%, Score 113.5, DB 2, Length 330;

Matches 38; Conservative 17; Mismatches 62; Indels 21; Gaps 5;

Db

16 AMKSLQDQNRKRYLAQTDTDGV-----GEETCVSVTAERIKKRLNATITLKN 68

114 AM-----SNDGYLATSR-DKSWIMETDESGEYICISVLQEH--SDQVKHYIMPS 164

69 KHLTDLKESHTITWKAVDYTT-----NRIKYETGRTQTEDVEVFSDYKNCVIF 123

165 EALASSSYDDTVRIWKYDDDMECVAVLNHGEYVMSDDPKTEGVRLCSGSDSYTR 224

124 VPKERGSDEQYELWVSE 141

225 VMKYMGEDEDDQOEYVCE 242

Db

RESULT 2

AB1247

two-component sensor histidine kinase 11sk [imported] - Listeria monocytogenes (strat

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AB1247

R:Glaser, P.; Frangou, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entlan, K.D.; Fshih,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.;

ok, C.; Schuete, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland  
 A>Title: Comparative genomics of *Listeria* species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AB1247  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-483 <GLA>  
 A:Cross-references: GB:NC\_003210; PIDN:CAC9456.1; PID:gl6410794; GSPDB:GN00177  
 A:Experimental source: strain Ecd-e  
 C:Genetics:  
 A:Gene: llsk

```

Query Match      8.8%; Score 87; DB 2; Length 483;
Best Local Similarity 24.0%; Pred. No. 5.7;
Matches 37; Conservative 29; Mismatches 58; Indels 30; Gaps 7;

QY 18 KSLQDDONKRRYYLAQATQTDTGVDWGEETFCVSAETKIGKKKLATILYKRNKHTLDLKE 77
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 308 KKLVGEMDLDSRAPDLSIQTK-----LQTTVDNAIVEGY---RRNPEVYAEVFTT-LKED 359
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 78 HETTVKAKDY-----TTENGKIKETQGTPTQTEEDVFSVDYKNCVDYIWPVKRGS 130
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 360 DTDDRALLQNHLEQLILLIIMDNAVKYSGDTEV---DMHYEKQKQIHI----- 406
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 131 DEGDYELWVSDDKIDKIPDCKCFMNAVFAOOQK 164
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 407 DVRDYGGISQSEELDKIFN--RTRYVDKAKRSEK 438
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 3  
AF1609  
two-component sensor histidine kinase lisk [imported] - *Listeria innocua* (strain Cl1p11262)  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence.revision 27-Nov-2001 #text.change 27-Nov-2001  
C:Accession: AF1609  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Esnlt, H. D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madheno, E.; Maitournam, A.; Meok, C.; Schlueder, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A. Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1609  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-483 <G1A>  
A:Cross-references: GB:AL592022; PIDN:CAC96646.1; PID:g16413888; GSPDB:GN00178  
A:Experimental source: strain Cl1p11262  
C:Genetics:  
;Gene: lisk

```

Query March      8.8%; Score 87; DB 2; Length 483;
Best Local Similarity 24.0%; Pred. No. 5.7;
Matches 37; Conservative 29; Mismatches 58; Indels 30; Gaps 7;

QY      18 KSLDODONKRYIAQAQTOTTDGVGWEEFTCVSTAERIGKKILNATLLYNKHILTDLKES 77
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       308 KLIWEOMLDSRAEQISOTKE----LQITDVNATVEGV---RRNFENVMEFETFL-KLED 359
          :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY       78 HETTYWKADY-----TTENGIKYETGTGRQTLEDFVFSVDYNCNDVIYVPKERGS 130
          :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

Db        360 DTDFRALIQQNHLEQLIIIMDNNAVVKYSGGTGV---DMHVYEKKQIH----- 406
          :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY       131 DEGDYELWASEDKTLIPDCCKFTMAFAOOOEK 164
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Db        407 DVRDYGEGISQEEDTKEN--RFRVDKASREK 438
          :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

RESULT 4  
S18453  
variant surface glycoprotein MtrA 1.6 - Trypanosoma brucei  
C:Species: Trypanosoma brucei

C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #tex\_change 20-Aug-1999  
C:Accession: S18453  
R:Carrington, M.; Miller, N.; Plum, M.; Roditt, I.; Wiley, D.; Turner, M.  
J. Mol. Biol. 221, 832-835, 1991  
A:Title: Variant specific glycoprotein of Trypanosoma brucei consists of two domains  
A:Reference number: S18445; MUID:92046037; PMID:1942032  
A:Accession: S18453  
A:Molecule type: mRNA  
A:Residues: 1-529 <CAR>  
A:Cross-references: EMBL:X56764; NID:g10474; PTDN:CAA40083.1; PID:g10475  
C:Superfamily: variant surface glycoprotein  
C:Keywords: glycoprotein; membrane protein

[illegible]

RESULT 5  
T03284  
myoblast c1ty protein - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 17-Nov-2000  
C:Accession: T03284  
R:Erickson, M.R.S.; Galletta, B.J.; Abmayr, S.M.  
J. Cell Biol. 138, 589-603, 1997  
A:Title: *Drosophila* myoblast c1ty encodes a conserved protein that is essential for m  
A:Reference number: Z14881; MUID:97392689; PMID:9245768  
A:Accession: T03284  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1970 <ERI>  
A:Cross-references: EMBL:AF007805; NID:g2367443; PID:g2367444  
C:Genetics:  
A:Gene: mbc  
A:Cross-references: FlyBase:FBgn0015513  
A:Map position: 3

Query Match	8.6%	Score 85.5:	DB 2:	Length 1970:
Best Local Similarity	21.0%	Pred. No. 45:		
Matches	35:	Conservative	27:	Mismatches 54: Indels 51: Gaps 6:
QY	44	EFTCVSYAERIGKKKLNATLLYKKKH-----LTDLKESHETIYVMKRAYDYTTENGKRY	97	
DB	455	EFAIATITSE-----KNVEASYCVANEGGYIMAPGVISTAGHQIPIDERYSSVYYYHDDKPRW	510	
QY	98	ETGCTRTQTEDEVFVSDYKNCQDYLVPKRGSGDEGDELYMWSBDKIDKIPDCKCFYMAV	157	
DB	511	Q-----ETFKIHVPIEDFKOCHLRFVLKHHSSNE-----QKDRTEK-----PFGLAY	552	
QY	158	FAQQQ-----EKTVRNRYTYDSSCKPAAQ	181	
DB	553	VRLMQANGTTTGGOHLLAYVKIDHKKYDKTIVANCYLELPATVAVELQ	599	

RESULT  
AG1085

hypothetical protein lmo0086 [imported] - *Listeria monocytogenes* (strain EGD-e)  
 C:Species: *Listeria monocytogenes*  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 A:Accession: AG1085  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgeir, O.; Entlan, K.D.; Fsihl, H.; Jones, L.M.; Karel, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maltournam, A.; Meok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AG1085  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1959 <GLA>  
 A:Cross-references: GB:NC\_003210; PIDN:CAC98301.1; PID:g16409445; GSPDB:GN00177  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: lmo0086

Query Match 8.5%; Score 85; DB 2; Length 1959;  
 Best Local Similarity 24.8%; Pred. No. 49;  
 Matches 35; Conservative 34; Mismatches 44; Indels 28; Gaps 7;

OY 49 SVTAETIG--KKKNATILYKKNHLDKESHER--ITVWKAYDYTTENGKYEYTOGTR 103  
 DB 1324 AISENIOYVDRKVLVDIGEDKONTYKANKNELFNVYKADOTITLVNSIRVGLPTR 1383  
 OY 104 TOTEDVFEVFSQYKNCVIVPKERGSDEGDELYVSEBDKIDKIPDCKFTMAYFAQ--Q 161  
 DB 1384 DVSTETELSPDIDY---VVKV-----EGKYDLIDGNGEKDY-----YXSEIK 1423  
 OY 162 QEKTVRNV-YTDSCKPAPAQ 181  
 DB 1424 TEKSLPSMTSTDYSPVAYGQ 1444

RESULT 7  
 706491  
 beta-fructofuranosidase (EC 3.2.1.26) - garden pea  
 N:Alternate names: Invertase  
 C:Species: *Pisum sativum* (garden pea)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-May-1999  
 C:Accession: T06491  
 R:Zhang, L.  
 submitted to the EMBL Data Library, January 1996  
 A:Reference number: Z15715  
 A:Accession: T06491  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-555 <ZHA>  
 A:Cross-references: EMBL:X85327; NID:g1160487; PID:g773259  
 A:Experimental source: cv. Little Marvel  
 C:Genetics:  
 A:Note: bfruct  
 C:Superfamily: beta-fructofuranosidase  
 C:Keywords: glycosidase; hydrolase

Query Match 8.5%; Score 84; DB 2; Length 555;  
 Best Local Similarity 23.5%; Pred. No. 12;  
 Matches 38; Conservative 22; Mismatches 52; Indels 50; Gaps 8;

OY 2 PTWANEALGSGYODAMKSLQODQNKRYLAQTOTTDGVMGEFFTCVSTAEKIKKRLN 61  
 DB 174 PTIAKINSSSFRDPTTS-----WLGR-----DGFWR-----VLIGSKIDTKGM- 212  
 OY 62 ATILYKKNHLDLKE-----SHETITWKAYDY-----TTENG---IKYE 98  
 DB 213 -AIIYKSKNFVDWYAKHPLHSABCTGMCECPDFPVLDKNLRTGVDTSRKGDDVVRHV 271  
 OY 99 TQGTTRTQTFEDVFEVFSQYKNCVIVPKERGSDEG-----DY 135

DB 272 LKVSIDTKHHDHYLIGSYDVVKVFPENGEDNGFVLRYDY 313

RESULT 8  
 706380  
 beta-fructofuranosidase (EC 3.2.1.26) - garden pea  
 N:Alternate names: cell wall invertase  
 C:Species: *Pisum sativum* (garden pea)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 22-Jun-1999  
 A:Accession: T06380  
 R:Zhang, L.; Cohn, N.S.; Mitchell, J.P.  
 submitted to the EMBL Data Library, May 1998  
 A:Reference number: Z15639  
 A:Accession: T06380  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-582 <ZHA>  
 A:Cross-references: EMBL:AF063246; NID:g3152879; PIDN:AACT17166.1; PID:g3152880  
 A:Experimental source: cv. Little Marvel  
 C:Genetics:  
 A:Gene: bfruct  
 A:Function: hydrolyzes sucrose to glucose and fructose  
 C:Superfamily: beta-fructofuranosidase  
 C:Keywords: glycosidase; hydrolase

Query Match 8.5%; Score 84; DB 2; Length 582;  
 Best Local Similarity 23.5%; Pred. No. 13;  
 Matches 38; Conservative 22; Mismatches 52; Indels 50; Gaps 8;

OY 2 PTWANEALGSGYODAMKSLQODQNKRYLAQTOTTDGVMGEFFTCVSTAEKIKKRLN 61  
 DB 174 PTIAKINSSSFRDPTTS-----WLGR-----DGFWR-----VLIGSKIDTKGM- 212  
 OY 62 ATILYKKNHLDLKE-----SHETITWKAYDY-----TTENG---IKYE 98  
 DB 213 -AIIYKSKNFVDWYAKHPLHSABCTGMCECPDFPVLDKNLRTGVDTSRKGDDVVRHV 271  
 OY 99 TQGTTRTQTFEDVFEVFSQYKNCVIVPKERGSDEG-----DY 135  
 DB 272 LKVSIDTKHHDHYLIGSYDVVKVFPENGEDNGFVLRYDY 313

RESULT 9  
 T00494  
 RNA-directed RNA polymerase (EC 2.7.7.48) - white button mushroom virus 1  
 N:Alternate names: RNA-dependent RNA polymerase  
 C:Species: white button mushroom virus 1  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T00494  
 R:van der Lende, T.R.; Duitman, E.H.; Gunnell, M.G.W.; Yu, L.; Wessels, J.G.H.  
 virology 217, 88-96, 1996  
 A:Title: Functional analysis of dsRNAs (L1, L3, L5 and M2) associated with isometric 3  
 A:Reference number: Z14158; MUID:96177126; PMID:8599239  
 A:Accession: T00494  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: genomic RNA  
 A:Residues: 1-1078 <VAN>  
 A:Cross-references: EMBL:X94361; NID:g1217917; PIDN:CAAG4144.1; PID:g219181  
 A:Genetics:  
 A:Gene: L1  
 C:Keywords: nucleotidyltransferase

Query Match 8.4%; Score 83.5; DB 2; Length 1078;  
 Best Local Similarity 21.4%; Pred. No. 32;  
 Matches 39; Conservative 33; Mismatches 71; Indels 39; Gaps 9;

OY 16 AAKSLQODQNKRYLAQTOTTDGVMGEFFTCVSTAEKIKKRLNATILYKN----- 68  
 DB 53 ALSNINPDNAPFIVQCELAEL-GVQIANVLIVKGI-----IQGIDRYKSLSRECK 107  
 OY 69 --KHLTDKESHETITWKAYDY-----TTENGKYEYTOGTRTQTFEDVFEVFSQYKNCV 121

Db 108 TLKHVHVQNNHNLWKEK-YKTHRVVLSYNDIYHKVGLIYV---GCYKSSFGEDDV 163  
QY 122 IFVFKERGSDEGDEYELWSEDKIDKIPD-----CCKFTMAFAOQDEKTVRNVYTD 173  
Db 164 L-----SGIDERTYD---AREKVNESNLKESKHPATKAVCAKYLQONFOAAPDYANA 215  
QY 174 SC 175  
Db 216 SC 217

RESULT 10  
112821  
hypothetical protein yong - Bacillus subtilis phage SPBC2  
C:Species: Bacillus subtilis phage SPBC2  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 15-Oct-1999  
C:Accession: T12821; A69914  
R:Lazarovic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Manuel, C.; Karamata, D.  
submitted to the EMBL Data Library, August 1997  
A:Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 prophage  
A:Reference number: 217583  
A:Accession: T12821  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-306 <LNA>  
A:Cross-references: EMBL:AF020713; NID:g3023478; PID:g3025535; PIDN:AC13030.1  
R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Brusch, C.; Caldwell, B.; Capuano, C.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerston, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fujita, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Gall  
lech, J.; Harwood, C.R.; Hentut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo, M.F.  
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lander, J.; Lazarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogatawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69914  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-306 <KCN>  
A:Cross-references: GB:299115; GB:AL009126; NID:g2634478; PIDN:CAB14028.1; PID:e1183557;  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yong

Query Match 8.3%; Score 82.5; DB 2; Length 306;  
Best Local Similarity 22.8%; Pred. No. 8;  
Matches 37; Conservative 20; Mismatches 56; Indels 49; Gaps 6;

QY 48 VSTAEIKGKKLNATL---YNNKHTDLKESHETTVMKAYDYTTENGICKETGTRT 104  
Db 72 VSAVNESESKKSGAGNLTGYMKNGKDYKSK---TWADSEFDEEEOQEVYTTQLLA 126  
QY 105 QTEDEYVVFSDYKNCVDYIFVFKERGSDEGDEYELWSEDKIDKIPDCKFTMAV----- 157  
Db 127 KSTEDI---SKEDIDELMPWGRGLDVEY-ITWLONEYTD-----FNRRKEDSKGM 174  
QY 158 -----FAOQDEKTVRNVYTDSSCKP 177  
Db 175 ELLINEICLTRLDIRKRRENGEKYDQOQKTLQDLGLSSNNKP 216

RESULT 11  
G71616  
hypothetical protein PFB0375W - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C>Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000

C:Accession: G71616  
R:Gardner, M.J.; Tetteilin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600; MUID:99021743; PMID:9804551  
A:Accession: G71616  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1802 <GAR>  
A:Cross-references: GB:AE001390; GB:AE001362; NID:g3845164; PIDN:AC71864.1; PID:g384  
A:Experimental source: clone 307  
C:Genetics:  
A:Gene: PFB0375W

Query Match 8.3%; Score 82.5; DB 2; Length 1802;  
Best Local Similarity 25.9%; Pred. No. 74;  
Matches 30; Conservative 13; Mismatches 42; Indels 31; Gaps 4;

QY 61 NATLYKKNHLLTDKESHETTVMKAYDYTTENGICKETGTRT-----RTQTEDEYVVFSDY 116  
Db 21 NNNITGGYDPCPLKEIYTNVHYKSYHKKENIKYKAEKENDIDNNKKDOLFVYNDH 80  
QY 117 KNCDFIVFKERGSDEGDEYELWSEDKIDKIPDCKFTMAVFAOQDEKTVRNVYTD 172  
Db 81 K-----IDDKIRKRI-QCQKNCITY-----HDIEKNSYVD 109

RESULT 12  
146001  
C4b-binding protein alpha chain - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: I46001; S43190  
R:Hillarp, A.; Thern, A.; Dahlback, B.  
J. Immunol. 153, 4190-4199, 1994  
A:Title: Bovine C4b binding protein. Molecular cloning of the alpha- and beta-chains  
A:Reference number: I46001; MUID:95015909; PMID:7930621  
A:Accession: I46001  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-610 <HIL>  
A:Cross-references: EMBL:Z31693; NID:g469117; PIDN:CAA83498.1; PID:g469118  
C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology  
F:50-107/Domain: complement factor H repeat homology <FH2>  
F:112-169/Domain: complement factor H repeat homology <FH2>  
F:174-234/Domain: complement factor H repeat homology <FH3>  
F:239-284/Domain: complement factor H repeat homology <FH3>  
F:299-362/Domain: complement factor H repeat homology <FH3>  
F:366-425/Domain: complement factor H repeat homology <FH6>  
F:429-483/Domain: complement factor H repeat homology <FH6>  
F:487-541/Domain: complement factor H repeat homology <FH8>

Query Match 8.2%; Score 81; DB 1; Length 610;  
Best Local Similarity 23.7%; Pred. No. 26;  
Matches 32; Conservative 14; Mismatches 47; Indels 42; Gaps 5;

QY 78 HETITWKAVD-----YTTENGICKETGTRTQTEDEYVVFSDYKNCVDYIFVFKERGS 131  
Db 315 HQOTEDQVYDIDGIVLSKYKHFQKRETDPTVTQSNLEMSPIECKKVCCEPRLNN 374  
QY 132 EGDYEL-----WSEDKID-----KIPDC---CKFTMA 156  
Db 375 YGSITLHRRPSTSTHCYISGDKISYKCHSKYMFALCTKHTGWSPTPECRPDC-SP 433  
QY 157 YFAOQDEKTVRNVYTD 171  
Db 434 VIAHGQKVVSKFT 448

RESULT 13  
S30484  
pol polyprotein - human immunodeficiency virus type 2



C:Species: human immunodeficiency virus type 2, HIV-2  
C:Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 23-Mar-2001  
C:Accession: S30484  
R:Go, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
submitted to the EMBL Data Library, December 1992  
A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
A:Reference number: S30460  
A:Accession: S30484  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-656 <GAG>  
A:Cross-references: EMBL:M87114  
C:Superfamily: pol polyprotein

Query Match 8.2%; Score 81; DB 2; Length 656;

Best Local Similarity 23.0%; Pred. No. 28; Mismatches 50; Indels 66; Gaps 9;

Matches 42; Conservative 25; Mismatches 50; Indels 66; Gaps 9;

OY 15 DAKSLQDQNKRYLAQATQTTDGVWGEFTCVSTAEGIKKKLNATILYKKNHLDL 74  
DB 126 DAFSIPDEEFQYTA-----FTLPSVNADEGKR-----YLYK-----VL 162  
OY 75 KESHEITWKA-----DYTTENG-----KETGCTRTQTFEDVFVSDYKNC---VIF 123  
DB 163 PQG-----WKGSPALFQYTRNVLLEPRKANSVDTLIQYMDLILASDRDLEHDKYVL 216  
OY 124 VPKRGSDG-----DYELWVSEDKIDKIPDCKFTMAVFAQOQEKTV 166  
DB 217 QLEKLNLDLGLTPEEKQKQPPYHMGLEYELMPTKMKLQTE-----LPQREMDTV 267  
OY 167 RNW 169  
DB 268 NDI 270

#### RESULT 14

S73939  
Oligonucleotide transport ATP-binding protein oppf - Mycoplasma pneumoniae (strain ATCC 29311)  
N:Alternate names: hypothetical protein G07\_orf81  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 02-Feb-2001  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Plöckl, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:97103885; PMID:8948633  
A:Accession: S73939  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-851 <HIM>  
A:Cross-references: EMBL:AE000058; GB:U00089; NID:q1674291; PIDN:AAB96261.1; PID:q167431  
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Gene: oppf  
A:Genetic code: SGC3  
C:Superfamily: unassigned ATP-binding cassette proteins: ATP-binding cassette homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:31-764/Domain: ATP-binding cassette homology #status atypical <ABCL>  
F:48-55/Region: nucleotide-binding motif A (P-loop)

Query Match 8.2%; Score 81; DB 2; Length 851;

Best Local Similarity 23.7%; Pred. No. 39; Mismatches 60; Indels 56; Gaps 8;

Matches 44; Conservative 26; Mismatches 60; Indels 56; Gaps 8;

OY 3 TWANFAKLSYODA-WKSLQ-DQNKRYLAQATQTTDGVWGEFTCVSTAEGIKKK- 59  
DB 251 TFOQAKVSTOAFATWKLQOTKQNLKAYRAQ-----MAEELQNKPR 292  
OY 60 --LNAITLYKKNHLDLKEHSHETTYWKAAYDTTENGIKYETGCTRTQTFEDVFVFS--- 114  
DB 293 IYLAAML-----TTKNYIKDSRQ--NTQLTDVDFAFSYND 326

OY 115 --DYKNCDFVFPKRGSDGDEYELWSE--DKIDKIPDCKFTMAVFAQOQEKTVRN 170  
DB 327 MVDKKRLVVLSEYKALPYFYDWNQNMNDRDELINAVFPDLIDVYALNDFAVNE 386  
OY 171 TDSSCK 176  
DB 387 SDAKAE 392

#### RESULT 15

138614  
helicase II - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C:Accession: 138614  
R:Stayton, C.L.; Dabovic, B.; Gullano, M.; Geetz, J.; Broccoli, V.; Giovannazzi, S.; B  
Hum. Mol. Genet. 3, 1957-1964, 1994  
A>Title: Cloning and characterization of a new human Xq13 gene, encoding a putative h  
A:Reference number: 138614; MUID:95179111; PMID:7874112  
A:Accession: 138614  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1641 <RES>  
A:Cross-references: EMBL:U09820; NID:q606832; PIDN:AAC50069.1; PID:q606833  
C:Genetics:  
A:Gene: RAD54L

Query Match 8.2%; Score 81; DB 2; Length 1641;

Best Local Similarity 26.4%; Pred. No. 89; Mismatches 56; Indels 40; Gaps 8;

Matches 42; Conservative 21; Mismatches 56; Indels 40; Gaps 8;

OY 32 QATQTTDGV---WGEFTCVSTAEGIKKKLNATILYKKNHLDLKEHSHETTYWKAAYD 88  
DB 54 QASASTDGVDLKSKRGQSFSLVKKVATK-----EKSRL-----KTTCCKAYQD 100  
OY 89 YTTENGIRYETGCTRTQTFEDVFVSDYKNCDFVFPKRGSD-----EGDYELWVS 140  
DB 101 GLSDIAEKFLKKGDSDETSED-----DKQSKGTEKKKPSDFKKYIKMEQYE--SS 153  
OY 141 EDKIDKIP---DCKFTMAVFAQOQEKTVRNVTYDSSCK 176  
DB 154 SDGTEKLEPRREICHFPKGI-----KQIKNGTIDGERK 186

Search completed: August 1, 2003, 12:16:36  
Job time : 15.52 secs

THIS PAGE BLANK (USPTO)